

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 10:59:31 ; Search time 18 Seconds
(without alignments)
1532.810 Million cell updates/sec

Title: US-09-869-677a-2

Sequence: 1 SSTGAKTAKSDKLKLVAVATNS.....PGDSYRAMMKMNDKISEGL 287

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	287	100.0	289	2	lipoprotein mtsa, fimbrial adhesin f
2	33	11.5	239	2	hypothetical prote
3	27	9.4	309	2	hypothetical prote
4	27	9.4	309	2	hypothetical prote
5	20	7.0	309	2	adhesin B precursor
6	20	7.0	310	2	adhesin - Streptoc
7	19	6.6	313	2	hypothetical prote
8	9	3.1	310	2	adhesion binding p
9	9	3.1	310	2	histidine binding p
10	8	2.8	255	2	hypothetical prote
11	8	2.8	309	2	hypothetical prote
12	8	2.8	383	2	aliphatic nitrilas
13	8	2.8	428	2	NDP-sugar dehydrog
14	8	2.8	586	2	HOL1 protein - yea
15	7	2.4	60	2	degradative enzyme
16	7	2.4	90	2	hypothetical prote
17	7	2.4	90	2	hypothetical prote
18	7	2.4	112	2	hypothetical prote
19	7	2.4	135	2	sorbose-permease p
20	7	2.4	135	2	probable sorbose p
21	7	2.4	143	2	adrenodoxin homolo
22	7	2.4	155	2	exdb protein limpo
23	7	2.4	155	2	hypothetical prote
24	7	2.4	155	2	hypothetical prote
25	7	2.4	162	2	probable regulator
26	7	2.4	166	2	conserved hypotnet
27	7	2.4	172	2	hypothetical prote
28	7	2.4	192	2	colicin v producti
29	7	2.4	192	2	hypothetical prote

After 17.1g date

ALIGNMENTS

30	7	2.4	192	2	D72081	conserved hypotnet
31	7	2.4	201	1	GOECR2	L(+)-tartrate dehy
32	7	2.4	201	2	A91122	L-tartrate dehydra
33	7	2.4	201	2	H85966	L-tartrate dehydra
34	7	2.4	201	2	A12328	hypothetical prote
35	7	2.4	237	2	T05875	hypothetical prote
36	7	2.4	240	2	S46997	B-cell receptor as
37	7	2.4	251	2	T34656	hypothetical prote
38	7	2.4	254	1	PUBYS	cold shock protein
39	7	2.4	272	2	T25608	hypothetical prote
40	7	2.4	297	2	H84168	molibdenum cofacto
41	7	2.4	308	2	C75097	hypothetical prote
42	7	2.4	332	2	S44206	hypothetical prote
43	7	2.4	334	2	T15332	hypothetical prote
44	7	2.4	340	2	T33313	hypothetical prote
45	7	2.4	361	1	A61648	translation releas

ALIGNMENTS

RESULT 1
T48894
lipoprotein mtsa, metal binding [validated] - Streptococcus pyogenes (fragment)
N:Alternate names: metal-binding protein mtsa
C:Species: Streptococcus pyogenes
C:Date: 02-Jun-2000 #sequence, revision 02-Jun-2000 #text, change 02-Sep-2000
C:Accession: T48894
R:Januliczky, R.W.; Pallon, J.; Bork, L.
Mol. Microbiol. 34, 596-606 (1999)
A:Title: Identification and characterisation of a Streptococcus pyogenes ABC transporter
A:Reference number: 224991
A:Accession: T48894
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-289 <JAN>
A:Cross-references: EMBL:AF180521; PIDN:AAD56939.1
A:Experimental source: strain APl
C:Gene(s):
A:Gene: mtsA
A>Note: transcription unit consisting out of mtsa (lipoprotein), mtsB (ATP-binding protein)
C:Function:
A:Description: Involved in iron and zinc uptake [validated, MUID:20032372]
A>Note: protein has affinity for Zn2+, Fe2+ and Cu2+; Zn2+ and Cu2+ compete for the same site
C:Superfamily: adhesin B

Query Match 100.0%; Score 287; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 9.9e-287;
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	SSTGAKTAKSDKLKLVAVATNSIIADMTKAIAGDKIDLHSIVPGDPHREPLPEAEKTS	60
DB	1	SSTGAKTAKSDKLKLVAVATNSIIADMTKAIAGDKIDLHSIVPGDPHREPLPEAEKTS	60
QY	61	NADVFYNGINLDEGGQAFKLVNAOQTKNKDDFAVSDGIDVYILEGASEKGEDEPHA	120
DB	61	NADVFYNGINLDEGGQAFKLVNAOQTKNKDDFAVSDGIDVYILEGASEKGEDEPHA	120
QY	121	WLNLNGIISKNIAKQIADPNKKEYEKNLKYAVLEKLDKEAKSKEDAIENKKL	180
DB	121	WLNLNGIISKNIAKQIADPNKKEYEKNLKYAVLEKLDKEAKSKEDAIENKKL	180
QY	181	IVTSGCFRYSKANGVSATYWEINTEEBEGPDDISSLIEKLYIKRSALFVSSVRR	240
DB	181	IVTSGCFRYSKANGVSATYWEINTEEBEGPDDISSLIEKLYIKRSALFVSSVRR	240
QY	241	PMEYTKDSGIPYISEIFTDISTAKKGPDSYRAMMKMNDKISEGL	287
DB	241	PMEYTKDSGIPYISEIFTDISTAKKGPDSYRAMMKMNDKISEGL	287

RESULT 2
A37186

RESULT 2	A37186	
----------	--------	--

fimbrial adhesin fima precursor - Streptococcus parasanguinis
C:Species: Streptococcus parasanguinis
C>Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 17-Nov-2000
C:Accession: A37186; S61912
R:Fenno, J.C.; LeBlanc, D.J.; Fives-Taylor, P.
Infect. Immun. 57, 3527-3533, 1989
A:Title: Nucleotide sequence analysis of a type 1 fimbrial gene of Streptococcus sanguis
A:Reference number: A37186; MUID:95035427; PMID:257255
A:Accession: A37186
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-309 <FEN1>
A:Cross-references: GB:M26130; NID:9567768; PIDN:AAA53077.1; PID:9153834
A:Experimental source: strain FW213
A:Note: the source is designated as Streptococcus sanguis
R:Fenno, J.C.; Shaikh, A.; Spetator, G.; Fives-Taylor, P.
Mol. Microbiol. 15, 849-863, 1995
A:Title: The fima locus of Streptococcus parasanguis encodes an ATP-binding membrane protein
A:Reference number: S61910; MUID:95319327; PMID:7596287
A:Accession: S61912
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-309 <FEN2>
A:Cross-references: EMBL:M26130; NID:9567768; PIDN:AAA53077.1; PID:9153834
A:Experimental source: strain FW213
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
A:Note: this publication is not cited in GenBank entry SRRSRA, release 117.0
A:Note: the source is designated as Streptococcus parasanguis
C:Genetics:
A:Gene: fima
C:Superfamily: adhesin B
C:Keywords: blocked amino end; lipoprotein; membrane protein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-109/Product: fimbrial adhesin fima #status predicted <MAT>
F:21/Binding site: sn-2,3-diacylglycerol (cys) (covalent) #status predicted
F:21/Modified site: fatty acylated amino end (cys) (in mature form) #status predicted

Query Match 11.5%; Score 33; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.2e-25;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 181 IYVSEGCFFKFSKAYGVPASVIMEINTEEGCTP 213
Db 201 IYVSEGCFFKFSKAYGVPASVIMEINTEEGCTP 233

RESULT 3
Hypothetical protein SPI650 [Imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C:Accession: H95191
R:Nettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held
on, J.D.; Umeyam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Kadane, D.; Holtzapfel,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: H95191
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-309 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK75729.1; PID:914973140; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SPI650
C:Superfamily: adhesin B

Query Match 9.4%; Score 27; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.7e-19;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 IAGDKIDLHSIVPIGDPHEHYEPLPED 55
Db 49 IAGDKIDLHSIVPIGDPHEHYEPLPED 75

RESULT 4
E98058
Hypothetical protein psaa [Imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: E98058
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.
e, R.; LeBlanc, D.J.; Lee, L.N.; Letkowitz, E.J.; Lu, J.; Matsushima, P.; McAdams, S.
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: E98058
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-309 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAL00298.1; PID:915459154; GSPDB:GN00174
C:Genetics:
A:Gene: psaa
C:Superfamily: adhesin B

Query Match 9.4%; Score 27; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.7e-19;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 IAGDKIDLHSIVPIGDPHEHYEPLPED 55
Db 49 IAGDKIDLHSIVPIGDPHEHYEPLPED 75

RESULT 5
A43583
Adhesin B precursor - Streptococcus sanguis
C:Species: Streptococcus sanguis
C>Date: 12-Jan-1993 #sequence_revision 12-Jan-1993 #text_change 24-Sep-1999
C:Accession: A43583
R:Ganeshkumar, N.; Hannam, P.M.; Kolenbrander, P.E.; McBride, B.C.
Infect. Immun. 59, 1093-1099, 1991
A:Title: Nucleotide sequence of a gene coding for a saliva-binding protein (SsAB) firo
A:Reference number: A43583; MUID:91147187; PMID:1671775
A:Accession: A43583
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-309 <GAN>
A:Cross-references: GB:M63481; NID:9153825; PIDN:AAC98426.1; PID:9153826
C:Superfamily: adhesin B

Query Match 7.0%; Score 20; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 111 SEKKEDPHAMLNENKIIY 130
Db 131 SEKKEDPHAMLNENKIIY 150

RESULT 6
T1151
Adhesin - Streptococcus gordonii
C:Species: Streptococcus gordonii
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-Aug-1999
C:Accession: T1151
R:Kolenbrander, P.E.; Andersen, R.N.; Ganeshkumar, N.
Infect. Immun. 62, 4469-4480, 1994
A:Title: Nucleotide sequence of the Streptococcus gordonii PK48 coaggregation adhesi
A:Reference number: Z17283; MUID:95012638; PMID:7927711
A:Accession: T1151

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-310 <KOL>
A:Cross-references: EMBL:L11577; NID:g110629; PID:g110633
A:Experimental source: strain PK488
C:Genetics:
A:Gene: scaa
C:Superfamily: adhesin B

Query Match 7.0%; Score 20; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 VPSAYIWEINTEEGTPDQ1 216
|||||
DB 218 VPSAYIWEINTEEGTPDQ1 237

RESULT 7

hypothetical protein mtsA [imported] - Lactococcus lactis subsp. lactis (strain ILL403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: D86789
R:Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A:Reference number: AB6625; MUID:21235186; PMID:11337471
A:Accession: D86789
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-313 <STO>
A:Cross-references: GB:AA005176; PID:g12724295; PIDN:AAK05414.1; GSPDB:GN00146
A:Experimental source: strain ILL403
C:Genetics:
A:Gene: mtsA
C:Superfamily: adhesin B

Query Match 6.6%; Score 19; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 198 PSAYIWEINTEEGTPDQ1 216
|||||
DB 222 PSAYIWEINTEEGTPDQ1 240

RESULT 8

adhesion binding proteins and lipoproteins with multiple specificity for metal cations
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AG1305
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkyl, G.; Madueno, E.; Maitournam, A.; Me
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1305
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-310 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC99925.1; PID:g16411301; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo1847
C:Superfamily: adhesin B

Query Match 3.1%; Score 9; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 0.55;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 200 AYIWEINTE 208
|||||
DB 221 AYIWEINTE 229

RESULT 9

adhesion binding protein and lipoprotein with multiple specificity for metal cations
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AG1677
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkyl, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1677
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-310 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC97191.1; PID:g16414462; GSPDB:GN00178
A:Experimental source: strain C1p11262
C:Genetics:
A:Gene: lln1961
C:Superfamily: adhesin B

Query Match 3.1%; Score 9; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 200 AYIWEINTE 208
|||||
DB 221 AYIWEINTE 229

RESULT 10

histidine transport protein Hsp PA2926 [imported] - Pseudomonas aeruginosa (strain P
D83281
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: D83281
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
.; Lohy, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: D83281
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-255 <STO>
A:Cross-references: GB:AE004718; GB:AE004091; NID:g9949009; PIDN:AA606314.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: hsp; PA2926

Query Match 2.8%; Score 8; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 208 EEEGTPDQ 215
|||||
DB 229 EEEGTPDQ 236

RESULT 11

H9832
hypothetical protein SA0587 [imported] - Staphylococcus aureus (strain N315)

C:Species: *Staphylococcus aureus*
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: H89832
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: H89832
 A:Molecule type: DNA
 A:Status: preliminary
 A:Residues: 1-309 <KUP>
 A:Cross-references: GB:BA000018; PTD:913700522; PIDN:BA041819.1; GSPDB:GN00149
 A:Experimental source: Strain N315
 C:Genetics:
 A:Gene: SA0587
 C:Superfamily: adhesin B

Query Match 2.8%; Score 8; DB 2; Length 309;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 43 GDDPHEYE 50
 |||||
 DB 63 GDDPHEYE 70

RESULT 12
 A43470
 aliphatic nitrilase - *Rhodococcus rhodochrous*
 C:Species: *Rhodococcus rhodochrous*
 C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000
 C:Accession: A43470
 R:Kobayashi, M.; Yanaka, N.; Nagasawa, T.; Yamada, H.
 A:Title: Primary structure of an aliphatic nitrile-degrading enzyme, aliphatic nitrilase
 A:Reference number: A43470; MUID:93003039; PMID:1390687
 A:Contents: K22
 A:Accession: A43470
 A:Status: preliminary
 A:Molecule type: DNA; protein
 A:Residues: 1-383 <KOB>
 A:Cross-references: GB:D12583; NID:g216931; PIDN:BA02127.1; PID:g216932
 A:Note: sequence extracted from NCBI backbone (NCBI:114184, NCBI:114185)
 C:Superfamily: nitrilase

Query Match 2.8%; Score 8; DB 2; Length 383;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 50 EPLPDAE 57
 |||||
 DB 272 EPLPDAE 279

RESULT 13
 G69988
 NDP-sugar dehydrogenase homolog ytaC - *Bacillus subtilis*
 C:Species: *Bacillus subtilis*
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: G69988
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Allout, G.; Azevedo, V.; Berte, C.; Biron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gallet, J.; Hatwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kunitz, K.; Lapidus, A.; Lardinois, Y.; M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadleir, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serron,

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Teipstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Whittes, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Dancho, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: G69988
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-428 <KUN>
 A:Cross-references: GB:Z99119; GB:AL009126; NID:g2635411; PIDN:CAB15064.1; PID:g26355
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: ytaC
 C:Superfamily: GDPmannose dehydrogenase

Query Match 2.8%; Score 8; DB 2; Length 428;
 Best Local Similarity 100.0%; Pred. No. 7.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 133 NIAKOLIA 140
 |||||
 DB 126 NIAKOLIA 133

RESULT 14
 S63386
 HOLL protein - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: protein N3494; protein YNR055c
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 20-Jun-2000
 C:Accession: S63386; S63387; S61930
 R:Phl, T.M.
 A:Title: submitted to the Protein Sequence Database, April 1996
 A:Reference number: S63346
 A:Accession: S63386
 A:Molecule type: DNA
 A:Residues: 1-586 <POH>
 A:Cross-references: EMBL:Z71670; NID:g1302573; PIDN:CAA96336.1; PID:g1302574; MIPS:YN
 A:Experimental source: strain S288C
 R:Duesterhoft, A.; Floeth, M.; Fritz, C.; Heuss-Nelitzel, D.; Hilbert, H.; Moestl, D.
 A:Reference number: S62944
 A:Accession: S63387
 A:Molecule type: DNA
 A:Residues: 1-305 <DUE>
 A:Cross-references: EMBL:Z71670; MIPS:YNR055c
 A:Experimental source: strain S288C
 R:Wright, M.B.; Howell, E.A.; Gaber, R.F.
 A:Title: submitted to the EMBL Data Library, May 1995
 A:Description: Gain-of-function mutations in the HOLL gene of *Saccharomyces cerevisiae*
 A:Reference number: S61930
 A:Accession: S61930
 A:Molecule type: DNA
 A:Residues: 1-509; 'F', 511-586 <WRI>
 A:Cross-references: EMBL:U42348; NID:g825500; PIDN:AAH47713.1; PID:g825501
 C:Genetics:
 A:Gene: SGD:HOLL
 A:Cross-references: SGD:S0005338; MIPS:YNR055c
 A:Map position: 14R
 C:Superfamily: yeast HOLL protein
 C:Keywords: yeast HOLL protein
 F:108-124/Domain: transmembrane #status predicted <TM1>
 F:134-150/Domain: transmembrane #status predicted <TM2>
 F:193-209/Domain: transmembrane #status predicted <TM3>
 F:224-240/Domain: transmembrane #status predicted <TM4>
 F:411-427/Domain: transmembrane #status predicted <TM5>
 F:453-469/Domain: transmembrane #status predicted <TM6>
 F:545-561/Domain: transmembrane #status predicted <TM7>

Query Match 2.8%; Score 8; DB 2; Length 586;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 NLENGITY 130
 |||||
 Db 24 NLENGITY 31

RESULT 15

degradative enzyme production factor degr - *Bacillus subtilis*

C:Species: *Bacillus subtilis*

C>Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 15-Oct-1999

C:Accession: I39957; A69614

R:Nagami, Y.; Tanaka, T.

J. Bacteriol. 166, 20-28, 1986

A:Title: Molecular cloning and nucleotide sequence of a DNA fragment from *Bacillus natto*

A:Reference number: I39957; MUID:6168015; PMID:3062853

A:Accession: I39957

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-60 <RES>

A:Cross-references: GB:M12917; NID:g143354; PIDN:AAA22671.1; PID:g143355

R:Yang, M.; Shimotsu, H.; Ferrari, E.; Henner, D.J.

J. Bacteriol. 169, 434-437, 1987

A:Title: Characterization and mapping of the *Bacillus subtilis* ptrR gene.

A:Reference number: I39958; MUID:87083406; PMID:3098734

A:Accession: I39958

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-60 <RES>

A:Cross-references: GB:M15318; NID:g143356; PIDN:AAA22672.1; PID:g143357

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd

A.; Ehrlich, S.D.; Emmerston, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallet

lech, J.; Harwood, C.R.; Henaui, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

T.; Muthers, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: A69614

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-60 <KUR>

A:Cross-references: GB:299115; GB:AL009126; NID:g2634478; PIDN:CAB14112.1; PID:e1183641;

A:Experimental source: strain 168

C:Genetics:

A:Gene: degr

Query Match

Best Local Similarity 2.4%; Score 7; DB 2; Length 60;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 IAKGKGP 268

Db 27 IAKGKGP 33

RESULT 16

hypothetical protein AGR_C_3318 [imported] - *Agrobacterium tumefaciens* (strain C58, Cere

C:Species: *Agrobacterium tumefaciens*

C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: F97577

R:Goodner, B.; Hinkle, G.; Gatung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium*

A:Reference number: A97359; PMID:11743194

A:Accession: F97577

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-90 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK87575.1; PID:g15156915; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C_3318

A:Map position: circular chromosome

Query Match

Best Local Similarity 2.4%; Score 7; DB 2; Length 90;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 DVYLEG 109

Db 53 DVYLEG 59

RESULT 17

hypothetical protein Atu1805 [imported] - *Agrobacterium tumefaciens* (strain C58, Dupo

C:Species: *Agrobacterium tumefaciens*

C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: AE2798

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl

Kap, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AE2798

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-90 <KUR>

A:Cross-references: GB:AE006888; PIDN:AA142803.1; PID:g17740248; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu1805

A:Map position: circular chromosome

Query Match

Best Local Similarity 2.4%; Score 7; DB 2; Length 90;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 DVYLEG 109

Db 53 DVYLEG 59

RESULT 18

hypothetical protein RC0836 [imported] - *Rickettsia conorii* (strain Malish 7)

C:Species: *Rickettsia conorii*

C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001

C:Accession: D97804

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;

Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in *Rickettsia conorii* and *Rickettsia prowazekii*.

A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: D97804

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-112 <KUR>

A:Cross-references: GB:AE006914; PIDN:AAL03374.1; PID:g15619938; GSPDB:GN00173

C:Genetics:

A:Gene: RC0836

Query Match

Best Local Similarity 2.4%; Score 7; DB 2; Length 112;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 220 IEKLKVI 226
 |||||
 Db 104 IEKLKVI 110

RESULT 19

B91254
 sorbose-permease PTS system IIA component [Imported] - Escherichia coli (strain O157:H7,
 C:Species: Escherichia coli
 C>Date: 16-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C:Accession: B91254
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.-G.;
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: B91254
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-135 <NAV>
 A:Cross-references: GB:BA000007; PIDN:BAH38425.1; PID:G13364478; GSPDB:GN00154
 C:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: ECS002

Query Match 2.4%; Score 7; DB 2; Length 135;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 157 VAKLEKL 163
 |||||
 Db 44 VAKLEKL 50

RESULT 20

P86094
 probable sorbose PTS component Z5617 [Imported] - Escherichia coli (strain O157:H7, subs
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: P86094
 R:Perrine, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamocists, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: P86094
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-135 <STO>
 A:Cross-references: GB:AE005174; NID:G12518959; PIDN:AAG59218.1; GSPDB:GN00145; UWGP:Z56
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: Z5617

Query Match 2.4%; Score 7; DB 2; Length 135;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 157 VAKLEKL 163
 |||||
 Db 44 VAKLEKL 50

RESULT 21

A31574
 adrenodoxin homolog - chicken (fragment)
 N:Alternate names: testis steroidogenic ferredoxin precursor
 C:Species: Gallus gallus (chicken)
 C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 11-Jun-1999
 C:Accession: A31574
 R:Kagimoto, K.; McCarthy, J.L.; Waterman, M.R.; Kagimoto, M.

Biochem. Biophys. Res. Commun. 155, 379-383, 1988
 A:Title: Deduced amino acid sequence of mature chicken testis ferredoxin.
 A:Reference number: A31574; MUID:88326329; PMID:3415692
 A:Accession: A31574
 A:Molecule type: mRNA
 A:Residues: 1-143 <KRA>

A:Cross-references: GB:M21275; NID:G211098; PIDN:AAA48576.1; PID:G211099
 A:Note: authors isolated from cDNA library with a bovine adrenodoxin sequence probe
 C:Superfamily: ferredoxin [2Fe-2S]; ferredoxin [2Fe-2S] homology
 C:Keywords: 2Fe-2S; electron transfer; iron sulfur protein; metalloprotein; mitochond
 F;119/Domain: transit peptide (mitochondrion) (fragment) #status predicted <TNP>
 F;20-143/Product: adrenodoxin homolog #status predicted <NAV>
 F;48-112/Domain: ferredoxin [2Fe-2S] homology <FER>
 F;65,71,74,111/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Query Match 2.4%; Score 7; DB 2; Length 143;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 265 KGRPGDS 271
 |||||
 Db 41 KGRPGDS 47

RESULT 22

T44783
 exd protein [Imported] - Bordetella pertussis
 C:Species: Bordetella pertussis
 C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 17-Mar-2000
 C:Accession: T44783
 R:Pradel, E.; Guiso, N.; Loch, C.
 submitted to the EMBL Data Library, February 1999
 A:Description: Construction and characterization of Bordetella pertussis tonB mutants
 A:Reference number: Z22838
 A:Accession: T44783
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-155 <PRA>
 A:Cross-references: EMBL:AJ132741; PIDN:CAB53385.1
 A:Experimental source: strain Tohamat
 C:Genetics:
 A:Note: exd
 C:Superfamily: tolR protein

Query Match 2.4%; Score 7; DB 2; Length 155;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 27 KATAGDK 33
 |||||
 Db 92 KATAGDK 98

RESULT 23

T29947
 hypothetical protein T20D4.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
 C:Accession: T29947
 R:Minx, P.; Graves, T.
 submitted to the EMBL Data Library, November 1996
 A:Description: The sequence of C. elegans cofilin T20D4.
 A:Reference number: Z20712
 A:Accession: T29947
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-155 <MIN>
 A:Cross-references: EMBL:U80029; PIDN:AAB37590.1; GSPDB:GN00023; CESP:T20D4.7
 C:Experimental source: strain Bristol N2; clone T20D4
 C:Genetics:
 A:Gene: CESP:T20D4.7
 A:Map position: 5
 A:introns: 62/3; 108/2

C:Superfamily: Caenorhabditis elegans hypothetical protein C35B1.5

Query Match 2.4%; Score 7; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 159 KLEKLDK 165
|||||||
DB 22 KLEKLDK 28

RESULT 24
F97019

hypothetical protein CAC0969 [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: F97019
R:Moelling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C10
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: F97019
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-155 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK78945.1; PID:q15023875; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:

A:Gene: CAC0969

Query Match 2.4%; Score 7; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 160 LEKIDKE 166
|||||||
DB 32 LEKIDKE 38

RESULT 25
AH0674

probable regulatory protein SRY1514 [imported] - Salmonella enterica subsp. enterica ser
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AH0674
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AH0674
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-162 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01769.1; PID:q16502617; GSPDB:GN00176
C:Genetics:

A:Gene: SRY1514

Query Match 2.4%; Score 7; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 VAKLEKL 163
|||||||
DB 84 VAKLEKL 90

RESULT 26
AF2899

conserved hypothetical protein Atu2631 [imported] - Agrobacterium tumefaciens (strain

C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AF2899
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woe
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
: Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Rao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AF2899
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-166 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL43612.1; PID:q17741132; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:

A:Gene: Atu2631

Query Match 2.4%; Score 7; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 212 TPDQISS 218
|||||||
DB 48 TPDQISS 54

RESULT 27
G97674

hypothetical protein AGR_C_4771 [imported] - Agrobacterium tumefaciens (strain C58, C
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: G97674
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: G97674
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-172 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK88352.1; PID:q15157833; GSPDB:GN00169
C:Genetics:

A:Gene: AGR_C_4771

Query Match 2.4%; Score 7; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 212 TPDQISS 218
|||||||
DB 54 TPDQISS 60

RESULT 28
AI3437

colicin v production protein [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AI3437
R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
: Mazur, M.; Goltman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella meli
A:Reference number: AD3252; PMID:11756688
A:Accession: AI3437
A:Status: preliminary

A:Gene: preliminary

A:Molecule type: DNA
 A:Residues: 1-192 <RUR>
 A:Cross-references: GB:AE008917; PIDN:AA152668.1; PID:g17963493; GSPDB:GN00190
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BME11487
 A:Map position: 1

Query Match 2.4%; Score 7; DB 2; Length 192;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 218 VAKLEKL 224
 DB 162 SLIEKLK 168

RESULT 29
 H86543
 hypothetical protein CPJ0426 [imported] - Chlamydia pneumoniae (strain J138)
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C:Accession: H86543
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishitani, F.; Ouchi, K.; Shiba, T.; Ise, N.
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349; PMID:10871362
 A:Accession: H86543
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-192 <STO>
 A:Cross-references: GB:BA000008; NID:98978798; PIDN:BA98634.1; GSPDB:GN00142
 C:Genetics:
 A:Gene: CPJ0426

Query Match 2.4%; Score 7; DB 2; Length 192;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 157 VAKLEKL 163
 DB 126 VAKLEKL 132

RESULT 30
 D72081
 conserved hypothetical protein frameshifted CP0327 [imported] - Chlamydia pneumoniae
 N:Alternate names: hypothetical protein CT27 homolog
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
 C:Accession: D72081; G81589
 R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammell, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:99206606; PMID:10192388
 A:Accession: D72081
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-192 <ARN>
 A:Cross-references: GB:AE001625; GB:AE001363; NID:g4376695; PIDN:AD18570.1; PID:g437670
 A:Experimental source: strain Cw1029
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
 A:Reference number: AB1500; MUID:20150255; PMID:10684935
 A:Accession: G81589
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-192 <REA>
 A:Cross-references: GB:AE002195; GB:AE002161; NID:g7189246; PIDN:AA938182.1; PID:g718925
 A:Experimental source: strain AR39, HL cells
 C:Genetics:

A:Gene: CPN0426; CP0327
 Query Match 2.4%; Score 7; DB 2; Length 192;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 157 VAKLEKL 163
 DB 126 VAKLEKL 132

RESULT 31
 O0ECR2
 L(+)-tartarate dehydratase (EC 4.2.1.32), iron-dependent, beta chain - Escherichia coli
 C:Species: Escherichia coli
 C:Date: 30-Jun-1988 #sequence_revision 31-Oct-1997 #text_change 01-Mar-2002
 C:Accession: D65094; I70799; B29049
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: D65094
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-201 <BLAT>
 A:Cross-references: GB:AE000388; GB:U00096; NID:g1789441; PIDN:AA676098.1; PID:g17894
 A:Experimental source: strain K-12, substrain MG1655
 R:Keane, S.K.; Begg, C.; Bungard, S.; Guest, J.R.
 J. Gen. Microbiol. 139, 1523-1530, 1993
 A:Title: Identification of the L-tartarate dehydratase genes (tda and tdb) of Escherichia coli K-12.
 A:Reference number: I55714; MUID:93381464; PMID:8371115
 A:Accession: I70799
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-79; QYAYGK', 86-164, 'A', 166-201 <RES>
 A:Cross-references: GB:U14781; NID:g347083; PIDN:AA03062.1; PID:g347085
 R:Neslin, M.; Lupski, J.R.; Svec, P.; Godson, G.N.
 Gene 51, 149-161, 1987
 A:Title: Possible new genes as revealed by molecular analysis of a 5-kb Escherichia coli K-12 DNA fragment.
 A:Reference number: A91573; MUID:87248073; PMID:3297921
 A:Accession: B29049
 A:Molecule type: DNA
 A:Residues: 1-79; QYAYGK', 86-164, 'A', 166-201 <RES>
 A:Cross-references: GB:M6194; GB:X00773; NID:g147764; PIDN:AA72574.1; PID:g147766
 C:Genetics:
 A:Gene: ttdb
 A:Map position: 67 min
 C:Complex: a tetramer containing two pairs of non-identical chains, the alpha and beta
 C:Function:
 A:Description: catalyzes the stereospecific interconversion of oxaloacetate and (R,R) isopropylmalate in the presence of ferrous ion and thiol
 C:Superfamily: iron-dependent tartarate dehydratase beta chain; iron-dependent tartarate dehydratase
 C:Keywords: carbon-oxygen lyase; hydro-lyase
 F.13-177/Domain: iron-dependent tartarate dehydratase beta chain homology <TTDB>

Query Match 2.4%; Score 7; DB 1; Length 201;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 IAKNKKL 180
 DB 175 IAKNKKL 181

RESULT 32
 A91122
 L-tartarate dehydratase subunit B [imported] - Escherichia coli (strain 0157:H7, subst
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: A91122
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
 Gasawara, N.; Yasunaga, T.; Kubara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: A91122
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-201 <HAV>
 A:Cross-references: GB:BA000007; PIDN:BA37368.1; PID:q13363418; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: EC3945
 C:Superfamily: Iron-dependent tartrate dehydratase beta chain; Iron-dependent tartrate d

Query Match 2.4%; Score 7; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 IAEKKL 180
 |||||||
 Db 175 IAEKKL 181

RESULT 33
 H85966
 L-tartrate dehydratase, subunit B [imported] - *Escherichia coli* (strain O157:H7, substra
 C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85966
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 11er, L.; Grobeck, E.J.; Davis, N.W.; Llm, A.; Dinalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: H85966
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-201 <STO>
 A:Cross-references: GB:AE005174; NID:q12517641; PIDN:AGS8196.1; GSPDB:GN00145; UMG:244
 A:Experimental source: strain O157:H7, substrain EDJ933
 C:Genetics: ttdb
 A:Gene: ttdb
 C:Superfamily: Iron-dependent tartrate dehydratase beta chain; Iron-dependent tartrate c

Query Match 2.4%; Score 7; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 IAEKKL 180
 |||||||
 Db 175 IAEKKL 181

RESULT 34
 A12328
 Hypothetical protein all184 [imported] - *Nostoc* sp. (strain PCC 7120)
 C:Species: *Nostoc* sp.
 A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: A12328
 R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iritguchi
 Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: A12328
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-201 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BAV75883.1; PID:q17133319; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all184

Query Match 2.4%; Score 7; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 KLIYSE 185
 |||||||
 Db 74 KLIYSE 80

RESULT 35
 T05875
 Hypothetical protein T29A15.190 - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 13-Aug-1999
 C:Accession: T05875
 R:Revan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.;
 submitted to the Protein Sequence Database, March 1999
 A:Reference number: Z15455
 A:Accession: T05875
 A:Molecule type: DNA
 A:Residues: 1-237 <BEV>
 A:Cross-references: EMBL:AL035602
 A:Experimental source: cultivar Columbia; BAC clone T29A15
 C:Genetics:
 A:Map position: 4
 A:Introns: 68/3; 86/3; 117/3; 164/1; 198/2
 A:Note: T29A15.190

Query Match 2.4%; Score 7; DB 2; Length 237;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 162 KLDKEAK 168
 |||||||
 Db 168 KLDKEAK 174

RESULT 36
 S46997
 B-cell receptor associated protein BAP 29 - mouse
 C:Species: *Mus musculus* (house mouse)
 C:Date: 01-Feb-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
 C:Accession: S46997
 R:Kim, K.M.; Adachi, T.; Nielsen, P.J.; Terashima, M.; Lamers, M.C.; Koehler, G.; Ret
 EMBO J. 13, 3793-3800, 1994
 A:Title: Two new proteins preferentially associated with membrane immunoglobulin D.
 A:Reference number: S46997; MUID:94349927; PMID:8070407
 A:Accession: S46997
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-240 <KIM>
 A:Cross-references: GB:X78684; NID:9541729; PIDN:CA55351.1; PID:9541730

Query Match 2.4%; Score 7; DB 2; Length 240;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 175 AENKKLI 181
 |||||||
 Db 174 AENKKLI 180

RESULT 37
 T34656
 Hypothetical protein SC1A9.07 SC1A9.07 - *Streptomyces coelicolor*
 C:Species: *Streptomyces coelicolor*
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C:Accession: T34656
 R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream,
 submitted to the EMBL Data Library, December 1998
 A:Reference number: Z21552
 A:Accession: T34656
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-251 <SAU>
A:Cross-references: EMBL:AL034446; PIDN:CAA22377.1; GSPDB:GN00070; SCOEDB:SC1A9.07
A:Experimental source: strain A3(2)
A:Genetics: SCOEDB:SC1A9.07

Query Match 2.4%; Score 7; DB 2; Length 251;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 69 GINLEDC 75
DB 105 GINLEDC 111

RESULT 38
PUBVS

cold shock protein TIR1 - yeast (Saccharomyces cerevisiae)
N:Alternate names: 25k protein; protein YER011w; serine-rich protein SRP1
C:Species: Saccharomyces cerevisiae
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 23-Mar-2001
C:Accession: S05803; PMID:3139887
R:Marquet, D.; Guo, X.J.; Lauguin, G.J.M.
J. Mol. Biol. 202, 455-470, 1988
A:Title: Yeast gene SRP1 (serine-rich protein). Intragenic repeat structure and identifi
A:Reference number: S05803; MUID:89011972; PMID:3139887
A:Accession: S05803
A:Molecule type: DNA
A:Residues: 1-254 <MAR>
A:Cross-references: EMBL:X12775; NID:94539; PIDN:CAA31262.1; PID:94540
R:Districh, F.S.
Submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmid 9537, 9581, 9495, 9867, and lambda
A:Reference number: S50433
A:Accession: S50469
A:Molecule type: DNA
A:Residues: 1-254 <DIR>
A:Cross-references: EMBL:U18778; NID:9603592; PIDN:AAB64544.1; PID:9603603; GSPDB:GN0000
C:Genetics:
A:Gene: SGD:TIR1; SRP1; MIPS:YER011w
A:Cross-references: SGD:S0000813; MIPS:YER011w
A:Map position: 5R
C:Superfamily: serine-rich protein
C:Keywords: cell wall; tandem repeat; transmembrane protein
F:114-209/Region: 12-residue repeats
F:238-254/Domain: transmembrane #Status Predicted <TMM>

Query Match 2.4%; Score 7; DB 1; Length 254;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGTGAKT 7
DB 204 SGTGAKT 210

RESULT 39

T25608
hypothetical protein C33D12.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T25608
R:Martin, J.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid C33D12.
A:Reference number: Z20057
A:Accession: T25608
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-272 <MAR>
A:Cross-references: EMBL:U64600; PIDN:AAB04569.1; GSPDB:GN00028; CESP:C33D12.5
A:Experimental source: strain Bristol N2; clone C33D12

C:Genetics:
A:Gene: CESP:C33D12.5
A:Map position: X
A:Introns: 27/3; 82/3; 163/2; 240/1

Query Match 2.4%; Score 7; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 SDLKRV 16
DB 78 SDLKRV 84

RESULT 40

H84168
molybdenum cofactor biosynthesis protein [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: H84168
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky
; Lettshauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja
; Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.;
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: H84168
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-297 <STO>
A:Cross-references: GB:AE004437; NID:910579734; PIDN:AA018716.1; GSPDB:GN00138
C:Genetics:
A:Gene: moae

Query Match 2.4%; Score 7; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 97 AVSDGID 103
DB 265 AVSDGID 271

RESULT 41

C75097
hypothetical protein PAB1596 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: C75097
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
A:Reference number: A75001
A:Accession: C75097
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-308 <RAW>
A:Cross-references: GB:AJ248286; GB:AL096836; NID:95458366; PIDN:CAB50080.1; PID:9545
C:Genetics:
A:Gene: PAB1596

Query Match 2.4%; Score 7; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 LIEKRV 225
DB 263 LIEKRV 269

RESULT 42

S44206
 Hypothetical protein 437 - Coxiella burnetii
 C:Species: Coxiella burnetii
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Feb-1995
 C:Accession: S44206
 R:Williams, H.; Thiele, D.; Oswald, W.; Krauss, H.
 submitted to the EMBL Data Library, April 1994
 A:Reference number: S44206
 A:Accession: S44206
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-332 <WILL>
 A:Cross-references: EMBL:X78969

Query Match 2.4%; Score 7; DB 2; Length 332;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 218 SLIERLK 224
 |||||
 DB 312 SLIERLK 318

RESULT 43
 T15332
 Hypothetical protein B0336.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T15332
 R:Raich, A.
 submitted to the EMBL Data Library, July 1995
 A:Description: The sequence of C. elegans cosmid B0336.
 A:Reference number: Z18331
 A:Accession: T15332
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-334 <TAI>
 A:Cross-references: EMBL:U32305; NID:9912752; PID:9912757; PIDN:AAC46831.1; CESP:B0336.5
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Gene: CESP:B0336.5
 A:Introns: 19/3; 51/3; 160/1; 191/3; 291/2

Query Match 2.4%; Score 7; DB 2; Length 334;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 DKLKVA 17
 |||||
 DB 290 DKLKVA 296

RESULT 44
 T33313
 Hypothetical protein K02H11.6 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T33313
 R:Roehling, T.; Wohlmann, P.; Antoniou, B.
 submitted to the EMBL Data Library, May 1998
 A:Description: The sequence of C. elegans cosmid K02H11.
 A:Reference number: Z21320
 A:Accession: T33313
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-340 <ROH>
 A:Cross-references: EMBL:AF068720; PIDN:AAC17785.1; GSPDB:GN00023; CESP:K02H11.6
 A:Experimental source: strain Bristol N2; clone K02H11
 C:Genetics:
 A:Gene: CESP:K02H11.6
 A:Map position: 5
 A:Introns: 8/1; 32/1; 46/1; 76/1; 99/1; 196/1; 246/3; 292/2

Query Match 2.4%; Score 7; DB 2; Length 340;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 159 KLEKDK 165
 |||||
 DB 206 KLEKDK 212

RESULT 45
 A61648
 translation releasing factor RF-1 - Coxiella burnetii
 N:Alternate names: peptide chain release factor 1
 C:Species: Coxiella burnetii
 C:Date: 06-Jun-1997 #sequence_revision 06-Jun-1997 #text_change 16-Jul-1999
 C:Accession: A61648
 R:Williams, H.; Thiele, D.; Oswald, W.; Krauss, H.
 submitted to the EMBL Data Library, April 1994
 A:Reference number: S44206
 A:Accession: A61648
 A:Molecule type: DNA
 A:Residues: 1-361 <WILL>
 A:Cross-references: EMBL:X78969; NID:9623026; PIDN:CA55563.1; PID:9623028
 A:Experimental source: Isolate Nine Mile RSA493, phase I
 C:Genetics:
 A:Gene: pf1a; RF-1
 C:Function:
 A:Description: codon-specific peptide-chain-release factor responsible for release of
 A:Pathway: protein biosynthesis
 C:Superfamily: translation releasing factor
 C:Keywords: protein biosynthesis

Query Match 2.4%; Score 7; DB 1; Length 361;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 218 SLIERLK 224
 |||||
 DB 4 SLIERLK 10

Search completed: May 21, 2003, 11:13:19
 Job time : 22 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 11:11:18 ; Search time 29 Seconds
(without alignments)
291.185 Million cell updates/sec

Title: US-09-869-677a-2

Perfect score: 287

Sequence: 1 SSTGAKTKAKSDKLKVATNS.....PGDSYAMKMNLDKISEGL 287

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size: 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfill1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	9.8	310	1 US-07-791-377-2	Sequence 2, Appl1
2	28	9.8	310	4 US-08-356-106-2	Sequence 2, Appl1
3	27	9.4	289	4 US-08-961-083-20	Sequence 20, Appl1
4	27	9.4	309	2 US-08-715-131-2	Sequence 2, Appl1
5	27	9.4	309	4 US-09-221-753-2	Sequence 2, Appl1
6	20	7.0	293	4 US-09-071-035-496	Sequence 496, App
7	20	7.0	316	4 US-09-071-035-494	Sequence 494, App
8	11	3.8	308	4 US-09-071-035-28	Sequence 26, Appl
9	11	3.8	309	1 US-08-729-202-1	Sequence 1, Appl1
10	8	2.8	309	1 US-08-896-371-1	Sequence 1, Appl1
11	8	2.8	316	4 US-09-134-001C-5547	Sequence 5547, Ap
12	7	2.4	172	6 5242821-17	Sequence 211, App
13	7	2.4	207	4 US-07-199-637A-211	Sequence 6, Appl1
14	7	2.4	254	1 US-07-667-276A-6	Sequence 24, Appl
15	7	2.4	724	4 US-09-562-737-24	Sequence 3, Appl1
16	6	2.1	13	2 US-08-760-075A-3	Sequence 3, Appl1
17	6	2.1	13	4 US-08-338-546-3	Sequence 3, Appl1
18	6	2.1	13	4 US-09-659-084-3	Sequence 3, Appl1
19	6	2.1	17	4 US-08-602-999A-445	Sequence 445, App
20	6	2.1	17	4 US-09-500-124-445	Sequence 445, App
21	6	2.1	30	4 US-09-376-113-3	Sequence 3, Appl1
22	6	2.1	56	6 5217896-7	Sequence 3, Appl1
23	6	2.1	65	4 US-09-227-357-612	Sequence 3, Appl1
24	6	2.1	75	6 5320958-9	Sequence 2, Appl1
25	6	2.1	81	4 US-09-376-113-2	Sequence 2, Appl1
26	6	2.1	85	4 US-08-858-207A-318	Sequence 318, App

28	6	2.1	111	1 US-07-754-918A-11	Sequence 11, Appl
29	6	2.1	117	4 US-09-046-478-2	Sequence 2, Appl1
30	6	2.1	117	4 US-08-822-897C-2	Sequence 2, Appl1
31	6	2.1	117	4 US-09-608-810A-4	Sequence 4, Appl1
32	6	2.1	124	4 US-09-134-001C-5344	Sequence 5344, Ap
33	6	2.1	127	3 US-08-705-771-12	Sequence 12, Appl
34	6	2.1	141	2 US-08-411-726-5	Sequence 5, Appl1
35	6	2.1	141	6 5217896-3	Sequence 5, Appl1
36	6	2.1	155	4 US-09-615-192A-298	Sequence 298, App
37	6	2.1	161	4 US-08-858-207A-284	Sequence 284, App
38	6	2.1	168	4 US-09-376-113-5	Sequence 5, Appl1
39	6	2.1	172	2 US-08-923-738-2	Sequence 2, Appl1
40	6	2.1	172	2 US-08-923-738-4	Sequence 4, Appl1
41	6	2.1	172	4 US-08-936-165A-385	Sequence 385, App
42	6	2.1	173	1 US-08-193-977-10	Sequence 10, Appl
43	6	2.1	175	4 US-09-376-113-7	Sequence 7, Appl1
44	6	2.1	177	4 US-09-643-597-165	Sequence 165, App
45	6	2.1	177	4 US-09-643-597-166	Sequence 166, App

ALIGNMENTS

RESULT 1
US-07-791-377-2
; Sequence 2, Application US/0791377
; Patent No. 5422427
; GENERAL INFORMATION:
; APPLICANT: Russell, Harold
; APPLICANT: Tharpe, Jean A.
; APPLICANT: Sampson, Jacquelyn
; APPLICANT: O'Connor, Steven P.
; TITLE OF INVENTION: PNEUMOCOCCAL FIBRILL PROTEIN A
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARRY & CUSHMAN
; STREET: 1615 L. Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/791,377
; FILING DATE: 19911121
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/91969
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202)822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-791-377-2

Query Match 9.8%; Score 28; DB 1; Length 310;
Best local similarity 100.0%; Pred. No. 3.2e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 181 IYTSCECFYFSKAYGVPSAYIMEINTE 208
DB 202 IYTSCECFYFSKAYGVPSAYIMEINTE 229

RESULT 2
US-08-356-106-2
Sequence 2, Application US/08356106
Patent No. 6312944
GENERAL INFORMATION:
APPLICANT: Russell, Harold
APPLICANT: Tharpe, Jean A.
APPLICANT: Sampson, Jacquelyn
APPLICANT: O'Connor, Steven P.
TITLE OF INVENTION: PNEUMOCOCCAL FIMBRIAL PROTEIN A
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1615 L. Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,106
FILING DATE: 15-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/791,377
FILING DATE: 17-SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: Scott, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/91969
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-106-2

Query Match 9.8%; Score 28; DB 4; Length 310;
Best Local Similarity 100.0%; Pred. No. 3.2e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 181 IYVSEGCFFKFKYKAYGVPASVYWEINTE 208
Db 202 IYVSEGCFFKFKYKAYGVPASVYWEINTE 229

RESULT 3
US-08-961-083-20
Sequence 20, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 Inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-20

Query Match 9.4%; Score 27; DB 4; Length 289;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 IAGDKIDLHSIVPIGDPHEYPEPLPED 55
Db 29 IAGDKIDLHSIVPIGDPHEYPEPLPED 55

RESULT 4
US-08-715-131-2
Sequence 2, Application US/08715131
Patent No. 5854416
GENERAL INFORMATION:
APPLICANT: Sampson, Jacquelyn S.
APPLICANT: Russell, Harold
APPLICANT: Tharpe, Jean A.
APPLICANT: Ades, Edwin W.
APPLICANT: Carlone, George M.
TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE 37-kda SURFACE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Needle & Rosenberg, P.C.
STREET: 127 Peachtree Street, Suite 1200
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303-1811
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/715,131
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Spratt, Gwendolyn D.
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 14114.0200
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 688-0770
TELEFAX: (404) 688-9880
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 309 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-715-131-2

Query Match 9.4%; Score 27; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 2.9e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 IAGDKIDLHSIVPIGDPHEPEPLPED 55
|||||
Db 49 IAGDKIDLHSIVPIGDPHEPEPLPED 75

RESULT 5
US-09-221-753-2
Sequence 2, Application US/09221753
Patent No. 6217884
GENERAL INFORMATION:
APPLICANT: SAMPSON, JACQUELYN S.
APPLICANT: RUSSELL, HAROLD
APPLICANT: THARPE, JEAN A.
APPLICANT: ADES, EDWIN W.
APPLICANT: CARLONE, GEORGE M.
TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE 37. kda SURFACE
FILE REFERENCE: 64778 US
CURRENT APPLICATION NUMBER: US/09/221,753
EARLIER FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: US 07/791,377
EARLIER FILING DATE: 1991-09-17
EARLIER APPLICATION NUMBER: US 07/816,286
EARLIER FILING DATE: 1992-01-03
EARLIER APPLICATION NUMBER: US 08/222,179
EARLIER FILING DATE: 1994-04-04
EARLIER APPLICATION NUMBER: US 08/715,131
EARLIER FILING DATE: 1996-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 309
TYPE: PRT
ORGANISM: STREPTOCOCCUS PNEUMONIAE
US-09-221-753-2

Query Match 9.4%; Score 27; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 2.9e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 IAGDKIDLHSIVPIGDPHEPEPLPED 55
|||||
Db 49 IAGDKIDLHSIVPIGDPHEPEPLPED 75

RESULT 6
US-09-071-035-496
Sequence 496, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: GIL H. CHOI
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brooks
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 496:
SEQUENCE CHARACTERISTICS:
LENGTH: 293 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-496

Query Match 7.0%; Score 20; DB 4; Length 293;
Best Local Similarity 100.0%; Pred. No. 1.6e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 SEKGKEDPHAMVLENGIIT 130
|||||
Db 115 SEKGKEDPHAMVLENGIIT 134

RESULT 7
US-09-071-035-494
Sequence 494, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: GIL H. CHOI
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brooks
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 494:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-494

Query Match 7.0%; Score 20; DB 4; Length 316;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 111 SEKKEDPHAMNLENGIT 130
Db 138 SEKKEDPHAMNLENGIT 157

RESULT 8
US-09-071-035-28
Sequence 28, Application US/09071035
Patent No. 6448043

GENERAL INFORMATION:
APPLICANT: Gail H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8512
TELEFAX: (301) 309-8504
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-28

Query Match 3.8%; Score 11; DB 4; Length 289;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 DPHEYEPLPED 55
Db 44 DPHEYEPLPED 54

RESULT 9
US-09-071-035-26

Sequence 26, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gail H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-26

Query Match 3.8%; Score 11; DB 4; Length 308;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 DPHEYEPLPED 55
Db 63 DPHEYEPLPED 73

RESULT 10
US-08-729-202-1

Sequence 1, Application US/08729202
Patent No. 5700928
GENERAL INFORMATION:
APPLICANT: Hodgson, John
TITLE OF INVENTION: NOVEL SALIVA BINDING PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,202
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9521147.0
FILING DATE: 16-OCT-1995

APPLICATION NUMBER: 9604599.2
FILING DATE: 04-MAR-1996
APPLICATION NUMBER: 9616136.9
FILING DATE: 01-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31279
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 309 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
US-08-729-202-1

Query Match 2.8%; Score 8; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 GODPHEYE 50
|||||||
Db 63 GODPHEYE 70

RESULT 11
US-08-896-371-1
Sequence 1, Application US/08896371
Patent No. 5801234
GENERAL INFORMATION:
APPLICANT: Hodgson, John
APPLICANT: Burnham, Martin
TITLE OF INVENTION: NOVEL SALIVA BINDING PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,371
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/729,202
FILING DATE:
APPLICATION NUMBER: 9521147.0
FILING DATE: 16-OCT-1995
APPLICATION NUMBER: 9604599.2
FILING DATE: 04-MAR-1996
APPLICATION NUMBER: 9616136.9
FILING DATE: 01-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31279
TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 309 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
US-08-896-371-1

Query Match 2.8%; Score 8; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 GODPHEYE 50
|||||||
Db 63 GODPHEYE 70

RESULT 12
US-09-134-001C-5547
Sequence 5547, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5547
LENGTH: 316
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5547

Query Match 2.8%; Score 8; DB 4; Length 316;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GODPHEYE 50
|||||||
Db 70 GODPHEYE 77

RESULT 13
5242821-17
Patent No. 5242821
APPLICANT: PALVA, LIKKA, SIHAKOV, MERYI
TITLE OF INVENTION: LACTOCOCCUS PROMOTER AND SIGNAL
SEQUENCES FOR EXPRESSION IN BACTERIA
NUMBER OF SEQUENCES: 27
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/377,450
FILING DATE: 10-JUL-1989
SEQ ID NO: 17
LENGTH: 172
5242821-17

Query Match 2.4%; Score 7; DB 6; Length 172;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 EKNLKAY 156
DB 109 EKNLKAY 115

RESULT 14

US-09-199-637A-211
Sequence 211, Application US/09199637A
Patent No. 6355411
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick
APPLICANT: Goodman, Howard M.
APPLICANT: Rahme, Laurence G.
APPLICANT: Mahajan-Miklos, Shalina
APPLICANT: Tan, Man-Wah
APPLICANT: Cao, Hui
APPLICANT: Drenkard, Eliana
APPLICANT: Tsongalis, John
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
FILE REFERENCE: 00786/361002
CURRENT APPLICATION NUMBER: US/09/199,637A
CURRENT FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: 60/066,517
PRIOR FILING DATE: 1997-11-25
NUMBER OF SEQ ID NOS: 437
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 211
LENGTH: 207
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-211

Query Match
Best Local Similarity 2.4%; Score 7; DB 4; Length 207;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ATNSIIA 23
DB 149 ATNSIIA 155

RESULT 15

US-07-667-276A-6
Sequence 6, Application US/07667276A
Patent No. 5470971
GENERAL INFORMATION:
APPLICANT: Kondo, Keiji
APPLICANT: Inouye, Masayori
TITLE OF INVENTION: STRESS-INDUCED PROTEINS, GENES CODING
TITLE OF INVENTION: THEREFOR, TRANSFORMED CELLS OF ORGANISMS, METHODS AND
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weiser & Associates
STREET: 230 S. Fifteenth Street, Suite 500
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/667,276A
FILING DATE: 11-MAR-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763

REFERENCE/DOCKET NUMBER: 377.5351P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 254 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-07-667-276A-6

Query Match
Best Local Similarity 2.4%; Score 7; DB 1; Length 254;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSTGAKT 7
DB 204 SSTGAKT 210

RESULT 16

US-09-562-737-24
Sequence 24, Application US/09562737
Patent No. 6428967
GENERAL INFORMATION:
APPLICANT: Herz, Joachim
APPLICANT: Gotthardt, Michael
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/09/562,737
CURRENT FILING DATE: 2000-05-01
NUMBER OF SEQ ID NOS: 132
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 24
LENGTH: 724
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-24

Query Match
Best Local Similarity 2.4%; Score 7; DB 4; Length 724;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 LKVIKPS 229
DB 241 LKVIKPS 247

RESULT 17

US-08-760-075A-3
Sequence 3, Application US/08760075A
Patent No. 5942429
GENERAL INFORMATION:
APPLICANT: KIRSCHBAUM, Bernd
APPLICANT: MUELLNER, Stefan
APPLICANT: BARTLETT, Robert
TITLE OF INVENTION: NOVEL DEAH-BOX PROTEINS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,075A
FILING DATE: 04-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 19545126.0
FILING DATE: 04-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: GRANADOS, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 18748/309
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-760-075A-3

Query Match 2.1%; Score 6; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 KLIYTS 184
|||||
Db 1 KLIYTS 6

RESULT 18
US-09-338-546-3
Sequence 3, Application US/09338546
Patent No. 6251645
GENERAL INFORMATION:
APPLICANT: KIRSCHBAUM, Bernd
APPLICANT: MUELINER, Stefan
APPLICANT: BARTLETT, Robert
TITLE OF INVENTION: NOVEL DEAH-BOX PROTEINS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/338,546
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/760,075
FILING DATE: 04-DEC-1996
APPLICATION NUMBER: DE 19545126.0
FILING DATE: 04-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: GRANADOS, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 18748/309
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-338-546-3

Query Match 2.1%; Score 6; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 KLIYTS 184
|||||
Db 1 KLIYTS 6

RESULT 19
US-09-659-084-3
Sequence 3, Application US/09659084
Patent No. 6403299
GENERAL INFORMATION:
APPLICANT: KIRSCHBAUM, Bernd
APPLICANT: MUELINER, Stefan
APPLICANT: BARTLETT, Robert
TITLE OF INVENTION: NOVEL DEAH-BOX PROTEINS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/659,084
FILING DATE: 11-Sep-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/338,546
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: GRANADOS, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 18748/309
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-659-084-3

Query Match 2.1%; Score 6; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 KLIYTS 184
|||||
Db 1 KLIYTS 6

RESULT 20
US-08-602-999A-445

; Sequence 445, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A
; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 445:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-602-999A-445

Query Match 2.1%; Score 6; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 YEPLPE 54
DB 11 YEPLPE 16

RESULT 21
US-09-500-124-445
; Sequence 445, Application US/09500124
; Patent No. 6432920
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York

; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/500,124
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/602,999
; FILING DATE: 16-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 445:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-09-500-124-445

Query Match 2.1%; Score 6; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 YEPLPE 54
DB 11 YEPLPE 16

RESULT 22
US-09-376-113-3
; Sequence 3, Application US/09376113
; Patent No. 6451992
; GENERAL INFORMATION:
; APPLICANT: CUPP, Eddie Wayne
; APPLICANT: CUPP, Mary Smith
; TITLE OF INVENTION: Antithrombin Nucleotides and Proteins
; TITLE OF INVENTION: from Horn Fly
; FILE REFERENCE: 5721-10
; CURRENT APPLICATION NUMBER: US/09/376,113
; CURRENT FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Haematobia irritans
US-09-376-113-3

Query Match 2.1%; Score 6; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 DSGIPI 253
DB 16 DSGIPI 21

RESULT 23
5217896-7
; Patent No. 5217896
; APPLICANT: KRAMER, STEVEN P.; VALENZUELA, DAVID M.; REYNOLDS


```
;JR., FREDERICK H.;SORVILLO, JOHN M.
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES RECOGNIZING
; PARATHYROID HORMONE-LIKE PROTEIN
; NUMBER OF SEQUENCES: 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/292,263
; FILING DATE: 30-DEC-1988
; SEQ ID NO:7:
; LENGTH: 56
5217896-7

Query Match          2.1%; Score 6; DB 6; Length 56;
Best Local Similarity 100.0%; Pred. No.1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 264 KKGKPG 269
Db 5 KKGKPG 10

RESULT 24
US-09-227-357-612
; Sequence 612, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
```

```
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 612
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-227-357-612
```

```
Query Match          2.1%; Score 6; DB 4; Length 65;
Best Local Similarity 100.0%; Pred. No.1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 223 LKVIKP 228
Db 14 LKVIKP 19
```

```
RESULT 25
5320958-9
; Patent No. 5320958
; APPLICANT: INOUE, SUNIKO;HSU, MEI-YIN;EAGLE, SUSAN;
; INOUE, MASAYORI
; TITLE OF INVENTION: ISOLATED BACTERIAL REVERSE TRANSCRIPTASE
; NUMBER OF SEQUENCES: 24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/315,316
; FILING DATE: 24-FEB-1989
; SEQ ID NO:9:
; LENGTH: 75
5320958-9
```

```
Query Match          2.1%; Score 6; DB 6; Length 75;
Best Local Similarity 100.0%; Pred. No.1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 264 KKGKPG 269
Db 25 KKGKPG 30
```

```
RESULT 26
US-09-376-113-2
; Sequence 2, Application US/09376113
; Patent No. 6451992
; GENERAL INFORMATION:
; APPLICANT: Cupp, Eddie Wayne
; APPLICANT: Cupp, Mary Smith
; TITLE OF INVENTION: Antithrombin Nucleotides and Proteins
; TITLE OF INVENTION: from Horn Fly
; FILE REFERENCE: 5721-10
```

;; CURRENT APPLICATION NUMBER: US/09/376,113
;; CURRENT FILING DATE: 1999-08-17
;; NUMBER OF SEQ ID NOS: 7
;; SOFTWARE: FASTSEQ for Windows Version 3.0
;; SEQ ID NO 2
;; LENGTH: 81
;; TYPE: PRT
;; ORGANISM: Haematobia irritans
US-09-376-113-2

Query Match 2.1%; Score 6; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 DSGIP1 253
111111
DB 16 DSGIP1 21

RESULT 27
US-08-858-207A-318
; Sequence 318, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328e1 Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm1, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 318:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 85 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6348328e
US-08-858-207A-318

Query Match 2.1%; Score 6; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 GTPDOI 216
111111

DB 61 GTPDOI 66

RESULT 28
US-07-754-918A-11
; Sequence 11, Application US/07754918A
; Patent No. 5286484
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, R.S. et al
; TITLE OF INVENTION: NOCLEOTIDE SEQUENCE CODING FOR AN
; TITLE OF INVENTION: OUTER MEMBRANE PROTEIN FROM NEISSERIA MENINGITIDIS AND USE
; TITLE OF INVENTION: OF SAID PROTEIN IN VACCINE PREPARATIONS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stanger, Michaelson, Spivak and Wallace, Esq.
; STREET: Parkway 109 Office Center, 528 Newman Springs
; STREET: Road, P. O. Box 8489
; CITY: Red Bank
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5 1/4" 360Kb IBM compatible diskette
; COMPUTER: IBM PS/2 Model 80
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: Microsoft Word 5.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/754,918A
; FILING DATE: 19910905
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Michaelson, Peter L.
; REGISTRATION NUMBER: 30090
; REFERENCE/DOCKET NUMBER: Centro-2R
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)530-6671
; TELEFAX: (908)530-6584
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 111 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: p64k N-terminal comprising homology region with
; MOLECULE TYPE: "lysoil binding site" from E. coli Acetyl transferase
US-07-754-918A-11

Query Match 2.1%; Score 6; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 DKISEG 286
111111
DB 65 DKISEG 70

RESULT 29
US-09-046-479-2
; Sequence 2, Application US/09046479
; Patent No. 6291653
; GENERAL INFORMATION:
; APPLICANT: Shepherd, Paul O.
; APPLICANT: Delsher, Theresa A.
; TITLE OF INVENTION: MOTILIN HOMOLOGS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/046,479
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sawislak, Deborah A
REGISTRATION NUMBER: 37,438
REFERENCE/DOCKET NUMBER: 97-04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6672
TELEFAX: 206-442-6678
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-046-479-2

Query Match 2.1%; Score 6; DB 4; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 73 EDGQA 78
|||||
Db 60 EDGQA 65

RESULT 30
US-08-822-897C-2
Sequence 2, Application US/08822897C
Patent No. 6380158
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Deisher, Theresa A.
TITLE OF INVENTION: MOTILIN HOMOLOGS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,897C
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sawislak, Deborah A
REGISTRATION NUMBER: 37,438
REFERENCE/DOCKET NUMBER: 97-04
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6672
TELEFAX: 206-442-6678

TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 117 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-822-897C-2

Query Match 2.1%; Score 6; DB 4; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 73 EDGQA 78
|||||
Db 60 EDGQA 65

RESULT 31
US-09-608-810A-4
Sequence 4, Application US/09608810A
Patent No. 6420521
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Jaspers, Stephen R.
APPLICANT: Deisher, Theresa A.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: SGIP PEPTIDES
FILE REFERENCE: 99-51
CURRENT APPLICATION NUMBER: US/09/608,810A
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/141,592
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 117
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(23)
US-09-608-810A-4

Query Match 2.1%; Score 6; DB 4; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 73 EDGQA 78
|||||
Db 60 EDGQA 65

RESULT 32
US-09-134-001C-5344
Sequence 5344, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5344
LENGTH: 124

TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5344

Query Match
Best Local Similarity 2.1%; Score 6; DB 4; Length 124;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 KLEKLD 164
DB 33 KLEKLD 38

RESULT 33
US-08-705-771-12
Sequence 12, Application US/08705771
Patent No. 6054289
GENERAL INFORMATION:
APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,
TITLE OF INVENTION: Human Genes, Sequences and
TITLE OF INVENTION: Expression Products
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARILLA, BYRNE, BAIN, GILFILLAN,
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08705,771
FILING DATE: August 30, 1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-346 (PFI96)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 973-994-1700
TELEFAX: 973-994-1744
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 127 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-705-771-12

Query Match
Best Local Similarity 2.1%; Score 6; DB 3; Length 127;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 STAKG 266
DB 9 STAKG 14

RESULT 34
US-08-411-726-5
Sequence 5, Application US/08411726
Patent No. 5880093
GENERAL INFORMATION:
APPLICANT: BAGNOLI, FRANCO
TITLE OF INVENTION: Use of Parathormone, Its Biologically
TITLE OF INVENTION: Active Fragments and Correlated Peptides, for The Preparation
TITLE OF INVENTION: Pharmaceutical Compositions Useful for The Treatment of Pregn

NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: 1 Broadway
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10004
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
SOFTWARE: WordPerfect 6.1 for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/411,726
FILING DATE: 05-APR-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/02755
FILING DATE: 08-OCT-1993
APPLICATION NUMBER: MI-92A002331
FILING DATE: 09-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: PALMSE, Maria Luisa
REGISTRATION NUMBER: 34,402
REFERENCE/DOCKET NUMBER: 2111/1300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-425-7200
TELEFAX: 212-425-5288
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 141 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-411-726-5

Query Match
Best Local Similarity 2.1%; Score 6; DB 2; Length 141;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 KKGRPG 269
DB 90 KKGRPG 95

RESULT 35
5217896-3
Patent No. 5217896
APPLICANT: KRAMER, STEVEN P.; VALENZUELA, DAVID M.; REYNOLDS
J.R., FREDERICK H.; SORVILLO, JOHN M.
TITLE OF INVENTION: MONOCLONAL ANTIBODIES RECOGNIZING
PARATHYROID HORMONE-LIKE PROTEIN
NUMBER OF SEQUENCES: 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/292,263
FILING DATE: 30-DEC-1988
SEQ ID NO: 3:
LENGTH: 141
5217896-3

Query Match
Best Local Similarity 2.1%; Score 6; DB 6; Length 141;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 KKGRPG 269
DB 90 KKGRPG 95

RESULT 36
US-09-615-192A-298

```
; Sequence 298, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 298
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-615-192A-298

Query Match          2.1%; Score 6; DB 4; Length 155;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 TEEGCT 212
    |||||
Db 133 TEEGCT 138

RESULT 37
US-08-858-207A-284
; Sequence 284, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328e1 Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smitkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmil, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 284:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6348328e
US-08-858-207A-284

Query Match          2.1%; Score 6; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 ENKRLI 181
    |||||
Db 133 ENKRLI 138

RESULT 38
US-09-376-113-5
; Sequence 5, Application US/09376113
; Patent No. 6451992
; GENERAL INFORMATION:
; APPLICANT: Cupp, Eddie Wayne
; APPLICANT: Cupp, Mary Smith
; TITLE OF INVENTION: Antithrombin Nucleotides and Proteins
; FILE REFERENCE: 5721-10
; CURRENT APPLICATION NUMBER: US/09/376,113
; CURRENT FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Haematobia irritans
US-09-376-113-5

Query Match          2.1%; Score 6; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 DSGIRI 253
    |||||
Db 103 DSGIRI 108

RESULT 39
US-08-923-738-2
; Sequence 2, Application US/08923738A
; Patent No. 5948642
; GENERAL INFORMATION:
; APPLICANT: Burnham, Martin K. R.
; APPLICANT: Lonetto, Michael A.
; APPLICANT: Warren, Patrick V.
; TITLE OF INVENTION: No. 5948642e1 DNA Strand Resolution
; FILE REFERENCE: P50549-06
; CURRENT APPLICATION NUMBER: US/08/923,738A
; CURRENT FILING DATE: 1997-09-02
; EARLIER APPLICATION NUMBER: 60/027,032
; EARLIER FILING DATE: 1996-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-923-738-2

Query Match          2.1%; Score 6; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 IEKLV 225
```

DB 61 IERLKV 66

RESULT 40
US-08-923-738-4
Sequence 4, Application US/08923738A
Patent No. 5948642
GENERAL INFORMATION:
APPLICANT: Burnham, Martin K. R.
APPLICANT: Lonetto, Michael A.
APPLICANT: Warren, Patrick V.
TITLE OF INVENTION: No. 5948642el DNA Strand Resolution
FILE REFERENCE: P50549-06
CURRENT APPLICATION NUMBER: US/08/923,738A
CURRENT FILING DATE: 1997-09-02
EARLIER APPLICATION NUMBER: 60/027,032
EARLIER FILING DATE: 1996-09-24
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 172
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-923-738-4

Query Match 2.1%; Score 6; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 IERLKV 225
DB 61 IERLKV 66

RESULT 41
US-08-936-165A-385
Sequence 385, Application US/08936165A
Patent No. 6348582
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:

NAME: Gimm, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 385:
SEQUENCE CHARACTERISTICS:
LENGTH: 172 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-936-165A-385

Query Match 2.1%; Score 6; DB 4; Length 172;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 IERLKV 225
DB 64 IERLKV 69

RESULT 42
US-08-193-977-10
Sequence 10, Application US/08193977
Patent No. 5625031
GENERAL INFORMATION:
APPLICANT: WEBSTER, KEVIN R.
APPLICANT: COLEMAN, KEVIN G.
TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND
TITLE OF INVENTION: P34CDK2 CELL CYCLE REGULATORY KINASES AND HUMAN
TITLE OF INVENTION: PAPILLOMAVIRUS ET ONCOPROTEIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/193,977
FILING DATE: 08-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 5996-0016
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 173 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-193-977-10

Query Match 2.1%; Score 6; DB 1; Length 173;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 KISEGL 287

Db 152 KISECL 157

RESULT 43

US-09-376-113-7
 : Sequence 7, Application US/09376113
 : Patent No. 6451992
 : GENERAL INFORMATION:
 : APPLICANT: Cupp, Eddie Wayne
 : APPLICANT: Cupp, Mary Smith
 : TITLE OF INVENTION: Anticircumbin Nucleotides and Proteins
 : FILE REFERENCE: 5721-10
 : CURRENT APPLICATION NUMBER: US/09/376.113
 : CURRENT FILING DATE: 1999-08-17
 : NUMBER OF SEQ ID NOS: 7
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 7
 : LENGTH: 175
 : TYPE: PRT
 : ORGANISM: Haematobia irritans
 : US-09-376-113-7

Query Match 2.1%, Score 6; DB 4; Length 175;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 DSGIRI 253
 Db 110 DSGIRI 115

RESULT 44
 US-09-643-597-165
 : Sequence 165, Application US/09643597
 : Patent No. 6426072
 : GENERAL INFORMATION:
 : APPLICANT: Wang, Tongtong
 : APPLICANT: Fan, Liqun
 : APPLICANT: Kalos, Michael D.
 : APPLICANT: Bangur, Chaitanya S.
 : APPLICANT: Hosken, Nancy
 : APPLICANT: Fanger, Gary R.
 : APPLICANT: Li, Samuel X.
 : APPLICANT: Wang, Aijun
 : APPLICANT: Skeiky, Yasir A.W.
 : APPLICANT: Henderson, Robert A.
 : APPLICANT: McNeill, Patricia D.
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 : FILE REFERENCE: 210121.455C11
 : CURRENT APPLICATION NUMBER: US/09/643.597
 : CURRENT FILING DATE: 2000-08-21
 : NUMBER OF SEQ ID NOS: 369
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 165
 : LENGTH: 177
 : TYPE: PRT
 : ORGANISM: Homo sapien
 : US-09-643-597-165

Query Match 2.1%, Score 6; DB 4; Length 177;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 KKGKPG 269
 Db 126 KKGKPG 131

RESULT 45
 US-09-643-597-166

: Sequence 166, Application US/09643597
 : Patent No. 6426072
 : GENERAL INFORMATION:
 : APPLICANT: Wang, Tongtong
 : APPLICANT: Fan, Liqun
 : APPLICANT: Kalos, Michael D.
 : APPLICANT: Bangur, Chaitanya S.
 : APPLICANT: Hosken, Nancy
 : APPLICANT: Fanger, Gary R.
 : APPLICANT: Li, Samuel X.
 : APPLICANT: Wang, Aijun
 : APPLICANT: Skeiky, Yasir A.W.
 : APPLICANT: Henderson, Robert A.
 : APPLICANT: McNeill, Patricia D.
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 : FILE REFERENCE: 210121.455C11
 : CURRENT APPLICATION NUMBER: US/09/643.597
 : CURRENT FILING DATE: 2000-08-21
 : NUMBER OF SEQ ID NOS: 369
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 166
 : LENGTH: 177
 : TYPE: PRT
 : ORGANISM: Homo sapien
 : US-09-643-597-166

Query Match 2.1%, Score 6; DB 4; Length 177;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 KKGKPG 269
 Db 126 KKGKPG 131

Search completed: May 21, 2003, 11:13:54
 Job time : 31 secs

GenCore version 5.1.4.P5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 10:54:37 ; Search time 77 Seconds

(without alignments)
496.661 Million cell updates/sec

Title: US-09-869-677A-2

Perfect score: 1475
Sequence: 1 SSTGAKTKAKSKIKVAVATNS.....PGDSYAMKKNLKDLSGL 287

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.GeneSeq_101002:*

1: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*

2: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*

3: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*

4: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*

5: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*

6: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*

7: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*

8: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*

9: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*

10: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*

11: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*

12: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*

13: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*

14: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*

15: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*

16: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*

17: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*

18: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*

19: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*

20: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*

21: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*

22: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*

23: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1475	100.0	287	21 AAB07436	A streptococcal 11
2	1467	99.5	306	23 ABG66911	Streptococcus pyog
3	1401	95.0	274	23 ABP27001	Streptococcus poly
4	1253	84.9	308	23 ABP30114	Streptococcus poly
5	1179.5	80.0	309	20 AAY30350	37 kDa pneumococca
6	1179.5	80.0	309	20 AAM82496	S. pneumoniae 37-k
7	1179.5	80.0	309	22 AAU01906	Streptococcus pneu
8	1179.5	80.0	309	23 AAE19238	S. pneumoniae pneu
9	1174.5	79.6	289	19 AAW50722	Streptococcus pneu
10	1174.5	79.6	23	ABP54566	S. pneumoniae SP01

11	1174.5	79.6	310	21 AAY81668	Streptococcus pneu
12	1146.5	77.7	310	14 AAR37495	Pneumococcal fibr
13	1142	77.4	293	20 AAY00257	Enterococcus faeca
14	1142	77.4	293	23 AAB43476	E faecalis EF132 a
15	1142	77.4	316	20 AAY00256	Enterococcus faeca
16	1142	77.4	316	23 ABB43475	E faecalis EF132 p
17	1079.5	73.2	313	23 ABB54653	Lactococcus lactis
18	892	60.5	289	20 AAY00023	Enterococcus faeca
19	892	60.5	289	20 AAB43242	E faecalis EF008 a
20	892	60.5	308	20 AAY00022	Enterococcus faeca
21	892	60.5	308	23 AAB43241	E faecalis EF008 p
22	789	53.5	310	23 ABB49144	Listeria monocylog
23	755	51.2	168	23 ABB27434	Streptococcus poly
24	748.5	50.7	316	23 ABB40702	Staphylococcus epi
25	708	48.0	309	18 AAM26367	Staphylococcus aur
26	348	23.6	304	21 AAY74840	Neisseria meningit
27	342.5	23.2	280	23 ABB55469	Lactococcus lactis
28	341	23.1	308	21 AAY74841	Neisseria meningit
29	337	22.8	308	21 AAY74839	Neisseria gonorrhoe
30	328.5	22.3	313	23 ABB48436	Listeria monocylog
31	314	21.3	292	19 AAM54348	Streptococcal sal
32	313.5	21.3	506	23 ABB28185	Streptococcus poly
33	313.5	21.3	506	23 ABB29862	Streptococcus poly
34	307.5	20.8	318	16 AAR79722	ROM precursor TROM
35	307.5	20.8	318	18 AAM22134	Treponema pallidum
36	301.5	20.4	317	23 ABB49042	Listeria monocylog
37	273	18.5	515	23 ABB28186	Streptococcus poly
38	263	17.8	306	23 ABB26210	Streptococcus poly
39	262	17.8	344	23 ABB38679	Staphylococcus epi
40	260.5	17.7	318	22 AAG89778	C glutamylam prote
41	260.5	17.7	318	22 AAB76804	Corynebacterium q1
42	258.5	17.5	197	19 AAM55116	Streptococcus pneu
43	258.5	17.5	197	23 ABB54610	S. pneumoniae SP06
44	257	17.4	307	23 ABB26209	Streptococcus poly
45	254	17.2	307	23 ABB29704	Streptococcus poly

ALIGNMENTS

RESULT 1

AAB07436

AAB07436 standard; Protein; 287 AA.

AC AAB07436;

DT 20-OCT-2000 (first entry)

DE A streptococcal lipoprotein of an ABC transporter protein.

KW Lipoprotein; ATP-binding cassette transporter; MtsA; Immune response;

KW vaccine.

OS Streptococcus pyogenes.

PN WO200040729-A1.

PD 13-JUL-2000.

PE 30-DEC-1999; 99WO-CB04445.

PR 31-DEC-1998; 98GB-0028880.

PA (ACTI-) ACTINOVA LTD.

PI Bjoerck L, Janulczyk R;

PI WPI: 2000-465989/40.

DR N-PSTDB; AAM58808.

PT New polypeptide comprising an ATP-binding cassette transporter of

PT Streptococcus pyogenes for use as a vaccine against Streptococcus and

PT for assays that detect immune reactivity to the polypeptides in animals

PT and humans -
 XX
 PS Claim 1; Page 47-48; 55pp; English.
 XX
 CC The present sequence represents a Streptococcus pyogenes polypeptide
 CC which is a lipoprotein of an ATP-binding cassette transporter (Mtsa).
 CC The polypeptide generates a protective immune response to
 CC Streptococcus, preferably group A, such as S. pyogenes. It
 CC is used to prepare a vaccine against Streptococcus. The new
 CC polypeptides may be used in serological or cell mediated immune assays
 CC for the detection of immune reactivity to the polypeptides in animals
 CC and humans. Antibodies to Mtsa protein can be detected using an
 CC immunosay with the polypeptides.
 XX
 SQ Sequence 287 AA;
 Query Match 100.0%; Score 1475; DB 21; Length 287;
 Best Local Similarity 100.0%; Pred. No. 5.9e-115;
 Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSTGATKASDKLVVATNSIADMTKAIAGDKIDLSIPIGDPHEPEPLPDEAEKTS 60
 DB 1 SSTGATKASDKLVVATNSIADMTKAIAGDKIDLSIPIGDPHEPEPLPDEAEKTS 60
 QY 61 NADVIFNGINLEDDGQAMFTKLVKNAQTKNDYFAVSDGIDVITYLEGASEKGEDPRA 120
 DB 61 NADVIFNGINLEDDGQAMFTKLVKNAQTKNDYFAVSDGIDVITYLEGASEKGEDPRA 120
 QY 121 WLNLENGIITSKNIKAKLIADKPKNKETEKMLKAYVAKLEKIDKESKFDIAENKKL 180
 DB 121 WLNLENGIITSKNIKAKLIADKPKNKETEKMLKAYVAKLEKIDKESKFDIAENKKL 180
 QY 181 IYTSSECFYFSKAYGVSAYIWEINTEEGTDPDQISSLEKLYKPSALFEVSSVDRR 240
 DB 181 IYTSSECFYFSKAYGVSAYIWEINTEEGTDPDQISSLEKLYKPSALFEVSSVDRR 240
 QY 241 PMETVSKDSGIPYSEIFTDIATKAGKPGDSYAMKMWLNDKISEGL 287
 DB 241 PMETVSKDSGIPYSEIFTDIATKAGKPGDSYAMKMWLNDKISEGL 287
 RESULT 2
 ABG66911
 ID ABG66911 standard; Protein: 306 AA.
 AC ABG66911;
 XX
 DT 24-SEP-2002 (first entry)
 DE Streptococcus pyogenes strain M1 BVH-P3.
 XX
 KW Streptococcal polypeptide; gene therapy; pharyngitis; erysipelas;
 KW Impetigo; scarlet fever; invasive disease; bacteraemia;
 KW necrotising fasciitis; toxic shock; bacterial infection; vaccine.
 XX
 OS Streptococcus Pyogenes.
 XX
 PN MO200250107-A2.
 PD 24-JUN-2002.
 PF 21-DEC-2001; 2001MO-CA01853.
 PR 21-DEC-2000; 2000US-256940P.
 PA (SHIR-) SHIRE BIOCHEM INC.
 PI Martin D, Brodeur BR, Hamel J, Rioux S, Rheault P;
 DR WPI; 2002-500623/53.
 DR N-PSDB; ABR95351.
 PT Novel polypeptide useful for diagnostic Streptococcal bacterial

PT infection in a host susceptible to Streptococcal infection -
 XX
 PS Claim 1; Fig 4; 55pp; English.
 XX
 CC The invention describes an isolated Streptococcal polypeptide (I).
 CC (I) is useful for diagnosing streptococcal bacterial infection in a host
 CC susceptible to streptococcal infection by obtaining a biological sample
 CC from a host, incubating an antibody or its fragment reactive with (I)
 CC with the biological sample to form a mixture, and detecting specifically
 CC bound antibody or antigen or bound fragment in the mixture which
 CC indicates the presence of Streptococci or an antibody specific to
 CC Streptococci. A composition comprising (I) is useful for treatment (e.g.
 CC by gene therapy) of pharyngitis, erysipelas and impetigo, scarlet fever,
 CC and invasive diseases such as bacteraemia, necrotising fasciitis and
 CC toxic shock. The composition is also useful for therapeutic or
 CC prophylactic treatment of Streptococcus pyogenes bacterial infection in a
 CC host susceptible to Streptococcus pyogenes infection. This is the amino
 CC acid sequence of Streptococcus pyogenes BVH antigen used as a vaccine
 CC component for therapy and/or prophylaxis.
 XX
 SQ Sequence 306 AA;
 Query Match 99.5%; Score 1467; DB 23; Length 306;
 Best Local Similarity 99.3%; Pred. No. 3e-114;
 Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SSTGATKASDKLVVATNSIADMTKAIAGDKIDLSIPIGDPHEPEPLPDEAEKTS 60
 DB 18 SSTGATKASDKLVVATNSIADMTKAIAGDKIDLSIPIGDPHEPEPLPDEAEKTS 77
 QY 61 NADVIFNGINLEDDGQAMFTKLVKNAQTKNDYFAVSDGIDVITYLEGASEKGEDPRA 120
 DB 78 NADVIFNGINLEDDGQAMFTKLVKNAQTKNDYFAVSDGIDVITYLEGASEKGEDPRA 137
 QY 121 WLNLENGIITSKNIKAKLIADKPKNKETEKMLKAYVAKLEKIDKESKFDIAENKKL 180
 DB 138 WLNLENGIITSKNIKAKLIADKPKNKETEKMLKAYVAKLEKIDKESKFDIAENKKL 197
 QY 181 IYTSSECFYFSKAYGVSAYIWEINTEEGTDPDQISSLEKLYKPSALFEVSSVDRR 240
 DB 198 IYTSSECFYFSKAYGVSAYIWEINTEEGTDPDQISSLEKLYKPSALFEVSSVDRR 257
 QY 241 PMETVSKDSGIPYSEIFTDIATKAGKPGDSYAMKMWLNDKISEGL 287
 DB 258 PMETVSKDSGIPYSEIFTDIATKAGKPGDSYAMKMWLNDKISEGL 304

RESULT 3
 ABP27001
 ID ABP27001 standard; Protein: 274 AA.
 AC ABP27001;
 XX
 DT 02-JUL-2002 (first entry)
 DE Streptococcus polypeptide SEQ ID NO 3178.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 PN MO200234771-A2.
 PD 02-MAY-2002.
 PF 28-OCT-2001; 2001MO-GB04789.
 PR 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 PT

PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PI Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 DR WPI: 2002-352536/38.
 DR N-PSDB: ABN67632.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 PS
 PS Claim 1; Page 3471; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC Streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and anti-inflammatory
 CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX
 XX Sequence 274 AA:
 SO
 Query Match 95.0%; Score 1401; DB 23; Length 274;
 Best Local Similarity 99.6%; Pred. No. 8.3e-109;
 Matches 271; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 16 VATNSIINDMTKATAGDKIDHSIVPIGODPHEPELPEDAEKTSNADVIFFNGINLEDEG 75
 DB 1 VATNSIINDMTKATAGDKIDHSIVPIGODPHEPELPEDAEKTSNADVIFFNGINLEDEG 60
 QY 76 GOAMETKLVKNAOKTKNDYFAVSDGIDIVYLEGASEKGEKEDPHAMLNLENGIISKNA 135
 DB 61 GOAMETKLVKNAOKTKNDYFAVSDGIDIVYLEGASEKGEKEDPHAMLNLENGIISKNA 120
 QY 136 KOLIAKDPKNETYEKNIKAVAKLEKLDKEAKSKFDAIAENKKLIVTSEGCFFKFSKAY 195
 DB 121 KOLIAKDPKNETYEKNIKAVAKLEKLDKEAKSKFDAIAENKKLIVTSEGCFFKFSKAY 180
 QY 196 GVPASVYIWEINTEEGTDPDQISSLEKIKVIPSALFEVSSVDRRPMETVSKDSGIPITS 255
 DB 181 GVPASVYIWEINTEEGTDPDQISSLEKIKVIPSALFEVSSVDRRPMETVSKDSGIPITS 240
 QY 256 EIFTDSIAKKGKPGDSYAAKMKWNLDKISEGL 287
 DB 241 EIFTDSIAKKGKPGDSYAAKMKWNLDKISEGL 272
 RESULT 4
 ID ABP30114 standard; Protein: 308 AA.
 XX
 AC ABP30114;
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 9404.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus agalactiae.
 OS MO200234771-A2.
 XX
 XX 02-MAY-2002.
 XX
 XX 29-OCT-2001; 2001WO-GB04789.
 XX
 XX 27-OCT-2000; 2000GB-0026333.
 XX 24-NOV-2000; 2000GB-0028727.
 XX 07-NOV-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PI Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 DR WPI: 2002-352536/38.
 DR N-PSDB: ABN70745.
 XX
 XX
 XX New Streptococcus protein for the treatment or prevention of infection
 XX or disease caused by Streptococcus bacteria, such as meningitis, and
 XX for detecting a compound that binds to the protein -
 PS
 PS Claim 1; Page 4066; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC Streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and anti-inflammatory
 CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 CC
 SO Sequence 308 AA:
 Query Match 84.9%; Score 1253; DB 23; Length 308;
 Best Local Similarity 83.6%; Pred. No. 2.2e-96;
 Matches 234; Conservative 28; Mismatches 18; Indels 0; Gaps 0;
 QY 8 AKSDKIKVAVNSIINDMTKATAGDKIDHSIVPIGODPHEPELPEDAEKTSNADVIFF 67
 DB 27 AKNEKIKVAVNSIINDMTKATAGDKIDHSIVPIGODPHEPELPEDAEKTSNADVIFF 86
 QY 68 NGINLEDEGOAMETKLVKNAOKTKNDYFAVSDGIDIVYLEGASEKGEKEDPHAMLNLENG 127
 DB 87 NGINLEDEGOAMETKLVKNAOKTKNDYFAVSDGIDIVYLEGASEKGEKEDPHAMLNLENG 146
 QY 128 IYKSKNIAKOLIAKDPKNETYEKNIKAVAKLEKLDKEAKSKFDAIAENKKLIVTSEGC 187
 DB 147 IYKSKNIAKOLIAKDPKNETYEKNIKAVAKLEKLDKEAKSKFDAIAENKKLIVTSEGC 206
 QY 188 FFFESKAVGPSAVYIWEINTEEGTDPDQISSLEKIKVIPSALFEVSSVDRRPMETVSK 247
 DB 207 FFFESKAVGPSAVYIWEINTEEGTDPDQISSLEKIKVIPSALFEVSSVDRRPMETVSK 266
 QY 248 DSGIPITSIEFTDSIAKKGKPGDSYAAKMKWNLDKISEGL 287
 DB 267 DSGIPITSIEFTDSIAKKGKPGDSYAAKMKWNLDKISEGL 306

RESULT 5
 AAY30350
 ID AAY30350 standard; Protein; 309 AA.
 XX
 AC AAY30350;
 DT 09-NOV-1999 (first entry)
 XX
 DE 37 kDa pneumococcal surface adhesion A protein (PsaA).
 XX
 KW Pneumococcal surface adhesion A protein; PsaA; monoclonal antibody;
 XX vaccine; Streptococcus pneumoniae infection.
 XX Streptococcus pneumoniae.
 XX
 PN W09945121-A1.
 XX
 PD 10-SEP-1999.
 XX
 PF 26-FEB-1999; 99WO-US04326.
 XX
 PR 02-MAR-1998; 98US-0076565.
 XX
 PA (US\$) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Ades EW, Carlone GW, Sampson JS, Tharpe JA, Westerink MAJ;
 PI zeller JL;
 XX
 DR MPI: 1999-540849/45.
 DR N-PSDB: AA210411.
 XX
 PT New peptides corresponding to Streptococcus pneumoniae PsaA, used
 PT for treating or preventing Streptococcus pneumoniae infection in a
 PT subject
 XX
 PS Example 8; Page 53-54; 58pp; English.
 XX
 CC The present sequence represents a pneumococcal surface adhesion A
 CC protein (PsaA). The specification describes monoclonal antibodies which
 CC bind epitopes of the PsaA protein (see AAY30351-54). These peptides can
 CC be used in vaccines to prevent Streptococcus pneumoniae infections. The
 CC antibodies of the invention can also be used to detect S. pneumoniae in
 CC a sample or individual.
 XX
 Sequence 309 AA:

[illegible]

ID AAW62496 standard; Protein; 309 AA.
 XX AAW62496;
 AC
 XX
 DT 04-MAR-1999 (first entry)
 XX
 DE 5. pneumoniae 37-kDa surface adhesion A protein.
 XX
 XX Surface adhesion A protein; vaccine; detection; serotype; antibody;
 XX diagnostic; immunoassay; treatment; infection; anti-idiotype.
 KW
 XX Streptococcus pneumoniae.
 OS
 XX
 XX US5854416-A.
 PN
 XX
 PD 29-DEC-1998.
 XX
 XX
 PE 17-SEP-1996; 96US-0715131.
 XX
 PR 17-SEP-1996; 96US-0715131.
 PR 17-SEP-1991; 91US-0791377.
 PR 04-APR-1994; 94US-0222179.
 XX
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Ades EW, Carlone GW, Russell H, Sampson JS, Thairpe JA;
 DR WPI; 1999-095007/08.
 DR N-PSDB; AAV73914.
 XX
 XX Nucleic acid encoding the 37 kDa, surface adhesion A of Streptococcus
 PT pneumoniae - useful diagnostically and for production of
 XX recombinant polypeptides

PS Claim 1. Column 33-34; 20pp: English.

Query Match	80.0%:	Score 1179.5:	DB 20:	Length 309:			
Best Local Similarity	78.4%:	Pred. No. 2.9e-90:					
Matches 225:	Conservative 30:	Mismatches 31:	Indels 1:	Gaps 1:			
QY	2	STGAK-TAKSDKLT	LVATNSIADMTKATAGCKID	LHSIVP	IGODPHEPELPEDA	EKTS 60	
		:::::				:::::	
Db	21	ASGRKQPTTSGQRL	TKLVATNSIIDITKKN	IAGSKID	LHSIVP	IGODPHEPELPED	KAKTS 80
QY	61	NADYIFNGINLEDGQ	AMFTKLVKNAQCTKKKD	YFAVSDG	IDIVYILEGASEK	GKEDP	HA 120
		:::::				:::::	
Db	81	EADIDIFNGINLTG	GNAMFTKLVENAKT	ENKDDYFAVSD	VDVITYLEGON	EKGKEDP	HA 140
QY	121	WLNLENIIISKNIA	QOLIAKDPKNKETEY	EKRLKAYVAK	LEKIDKEAKS	FDMIAEN	KLU 180
		:::::				:::::	
Db	141	WLNLENIIIPAKK	IAQOLSAKDPNNKE	YEYENKLE	YIDDKID	KEKDFNNI	PAEKKL 200
QY	181	IYVSEGCFFKFSK	AVGPSAYIWEINTE	EEGTPDOJSL	ILEKTLKVP	SALEFESS	VDRR 240
		:::::				:::::	
Db	201	IYVSEGAFFKFSK	AVGPSAYIWEINTE	EEGTPDQITIV	EKLQRQVPS	LEFESS	VDRR 260
QY	241	PMETVSKDSG	IPYLSITFLDSIAK	GKRGDSYIAAMK	MNIDKTS	ESGL 287	
		:::::				:::::	

EVSSVDR 260
60 me
20

DB 25 TSSKSS-SDKLKVTTNSILADITKNAGDKIELHSIVPGODPHEPEPLPEDVKKTSQA 83

OY 63 DVIFNGINLEDDGQAMFTKLVKNAOKTKNNKDYFVSGIDIVYILEGASEGKEDPNAWL 122

DB 84 DLIFNGINLETTGNAMFTKLVKNAOKTKNNKDYFVSGIDIVYILEGQONQAGKEDPNAWL 143

OY 123 NLENNIIYSKINAKOLIAKDPKNEKYEKRLKAVAKLEKLDKRAKSDAIAENKTLIV 182

DB 144 NLENNIIYSKINAKOLIAKDPKNEKYEKRLKAVAKLEKLDKRAKSDAIAENKTLIV 203

OY 183 TSEGGFKYFSKAYGVPASAYIWEINTEEGPDDOISSILEKTKYKPSALFVSSVDRPM 242

DB 204 TSEGGFKYFSKAYGVPASAYIWEINTEEGPDDOISSILEKTKYKPSALFVSSVDRPM 263

OY 243 ETVSKDSGPIYSEIFETDSIAKKGKPGDSYAMKNNLDRKISEGL 287

DB 264 KTVSKDSNIPFAKILFTDSIAKEGEGDSYAMKNNLDRKISEGL 308

RESULT 13

AA00257

ID AAY00257 standard; Protein; 293 AA.

XX

AC AAY00257;

DT 20-APR-1999 (first entry)

XX

DE Enterococcus faecalis antigenic polypeptide fragment EF132.

XX

KW Enterococcus faecalis; infection; vaccine; immune response; diagnosis;

XX

OS detection; attenuation; antigenic.

XX

OS Enterococcus faecalis.

XX

PN WO980554-A2.

XX

PD 12-NOV-1998.

XX

PF 04-MAY-1998; 98WO-US08959.

XX

PR 14-NOV-1997; 97US-0066009.

PR 06-MAY-1997; 97US-0044031.

PR 16-MAY-1997; 97US-0046655.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

PI Bailey C, Choi GH, Hromocky J A, Kunsch CA;

PI WPI; 1999-070095/06.

DR N-PSDB; AAX20247.

DR

PT New isolated Enterococcus faecalis polynucleotides - used to develop

PT products for the detection of Enterococcus and for use in vaccines

PT for prevention or attenuation of Enterococcus infection

PT

PS Claim 9; Page 243; 301pp; English.

XX

XX

CC The present sequence represents an antigenic polypeptide fragment

CC isolated from Enterococcus faecalis. The present invention describes

CC genes, proteins and antigenic polypeptides isolated from E. faecalis.

CC The proteins can be used in vaccines for preventing or attenuating an

CC infection caused by a member of the Enterococcus genus in an animal.

CC They can also be used for detecting Enterococcus antibodies in a sample.

CC The nucleotide sequences can be used for detecting Enterococcus nucleic

CC acids. Products from the present invention can also be used for

CC screening compounds to identify agonists and antagonists of E. faecalis

CC protein activity.

XX

SQ Sequence 293 AA;

Query Match 77.4%; Score 1142; DB 20; Length 293;

Best Local Similarity 75.2%; Pred. No. 3.7e-87;

Matches 218; Conservative 28; Mismatches 40; Indels 4; Gaps 1;

OY 2 STGAK-----TAKSDKLKVATNSIADMTKRAIGDKIDLSHSIVPGODPHEPEPLPEDAE 57

DB 2 STNSKRDQDVASNEKIAKVVYVNSIADITENTAKOKIDLSHSIVPGODPHEPEPLPEDVQ 61

OY 58 KTSNADVIYINGINLEDDGQAMFTKLVKNAOKTKNNKDYFVSGIDIVYILEGASEGKED 117

DB 62 KTSKADLIFNGVNVNXTGQAMFTKLVKNAOKTKNNKDYFVSGIDIVYILEGASEGKED 121

OY 118 PHAMNLENGIYKINAKOLIAKDPKNEKYEKRLKAVAKLEKLDKRAKSDAIAEN 177

DB 122 PHAMNLENGIYKINAKOLIAKDPKNEKYEKRLKAVAKLEKLDKRAKSDAIAEN 181

OY 178 KRLIYTSBGCERYFSKAYGVPASAYIWEINTEEGPDDOISSILEKTKYKPSALFVSSV 237

DB 182 KRLIYTSBGCERYFSKAYGVPASAYIWEINTEEGPDDOISSILEKTKYKPSALFVSSV 241

OY 238 DRPMETVSKDSGPIYSEIFETDSIAKKGKPGDSYAMKNNLDRKISEGL 287

DB 242 DRPMETVSKDSGPIYSEIFETDSIAKKGKPGDSYAMKNNLDRKISEGL 291

RESULT 14

ABP43476

ID ABP43476 standard; Protein; 293 AA.

XX

AC ABP43476;

DT 05-AUG-2002 (first entry)

XX

DE E faecalis EF132 antigenic fragment.

XX

KW Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.

XX

OS Enterococcus faecalis.

XX

PN US2002045737-A1.

XX

PD 18-APR-2002.

XX

PF 04-MAY-1998; 98US-0071035.

XX

PR 04-MAY-1998; 98US-0071035.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

PI Choi GH, Bailey C, Hromocky J A, Kunsch CA;

PI WPI; 2002-425450/45.

DR N-PSDB; ABN98232.

DR

PT New genes and polypeptides from Enterococcus faecalis, useful as

PT vaccines for preventing, treating or attenuating an infection caused by

PT a member of the Enterococcus genus in an animal, particularly E.

PT faecalis

PT

PS Claim 9; Page 239; 255pp; English.

XX

XX

CC The present invention provides the protein and coding sequences of a

CC number of polypeptides from Enterococcus faecalis. The proteins can be

CC used as vaccines for preventing or attenuating an infection caused by a

CC member of the Enterococcus genus in an animal, particularly E. faecalis.

CC The polynucleotide is also useful for preventing or treating E. faecalis

CC infection. The present sequence is a protein of the invention.

XX

SQ Sequence 293 AA;

Query Match 77.4%; Score 1142; DB 23; Length 293;

Best Local Similarity 75.2%; Pred. No. 3.7e-87;

Matches 218; Conservative 28; Mismatches 40; Indels 4; Gaps 1;

OY 2 STGAK-----TAKSDKLKVATNSIADMTKRAIGDKIDLSHSIVPGODPHEPEPLPEDAE 57

DB 2 STNSKRDQDVASNEKIAKVVYVNSIADITENTAKOKIDLSHSIVPGODPHEPEPLPEDVQ 61

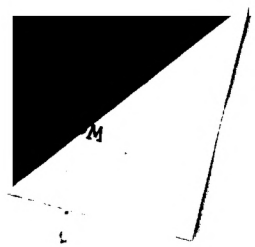
Db 2 STNSKDKTVAASNEKLVVATNSIIADITENIAKDKIDLSHIVPGDPHEPEPLPEDVQ 61
 QY 58 KTSNADVIIFYNGINLEDEGQAMFTKLVNNAOKTKNKKDYFAVSDGIDVITYLEGASERKED 117
 Db 62 KTSADLIFYNGVNLXGTGNMFTKLVXKXANKKEKNDFFASDGDIDVITYLEGOSERKED 121
 QY 118 PHAMINLENGIITYAKNIETKMLAEKDPDNKKFYKENDKYIEKIDSLDEAKSKFASIPND 177
 Db 122 PHAMINLENGIITYAKNIETKMLAEKDPDNKKFYKENDKYIEKIDSLDEAKSKFASIPND 181
 QY 178 KKLIVTSEGCERYFSKAYGVPASVAYIWEINTEEGTPDQISSLEIKLYIKPSALFEVSSV 237
 Db 182 KKLIVTSEGCERYFSKAYGVPASVAYIWEINTEEGTPDQIKHLVERLRTTKVPSLFVSSV 241
 QY 238 DRPMETVSKDSGIPYISEIFTDSIAKKGKPGDSYAMKKNLKDISEGL 287
 Db 242 DRPMETVSKDSGIPYISEIFTDSIAKKGKPGDSYAMKKNLKDISEGL 291

RESULT 15
 ID AAY00256 standard; Protein: 316 AA.
 AC AAY00256;
 XX 20-APR-1999 (first entry)
 DT 20-APR-1999 (first entry)
 XX Enterococcus faecalis protein EPI32.
 DE Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
 KW detection; attenuation; antigenic.
 KM Enterococcus faecalis.
 OS Enterococcus faecalis.
 XX W09850554-A2.
 PN 12-NOV-1998.
 PD 04-MAY-1998; 98WC-US08959.
 PF 14-NOV-1997; 97US-0066009.
 PR 06-MAY-1997; 97US-0044031.
 PR 16-MAY-1997; 97US-0046655.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Bailey C, Choi GH, Hromocky J A, Kunsch CA;
 PI WPI; 1999-070095/06.
 XX N-PSDB; AAX20246.
 DR New isolated Enterococcus faecalis polynucleotides - used to develop
 PT products for the detection of Enterococcus and for use in vaccines
 PT for prevention or attenuation of Enterococcus infection
 PS Claim 9; Page 243; 301pp; English.
 XX The present sequence represents a protein isolated from
 CC Enterococcus faecalis. The present invention describes genes, proteins
 CC and antigenic polypeptides isolated from E. faecalis. The proteins can
 CC be used in vaccines for preventing or attenuating an infection caused
 CC by a member of the Enterococcus genus in an animal. They can also be
 CC used for detecting Enterococcus antibodies in a sample. The nucleotide
 CC sequences can be used for detecting Enterococcus nucleic acids.
 CC products from the present invention can also be used for screening
 CC compounds to identify agonists and antagonists of E. faecalis protein
 CC activity.
 XX Sequence 316 AA;

Query Match 77.4%; Score 1142; DB 20; Length 316;
 Best Local Similarity 75.2%; Pred. NO. 4; 1e-87;
 Matches 218; Conservative 28; Mismatches 40; Indels 4; Gaps 1;

QY 2 STGAK----TAKSDKLVVATNSIIADITENIAKDKIDLSHIVPGDPHEPEPLPEDAE 57
 Db 25 STNSKDKTVAASNEKLVVATNSIIADITENIAKDKIDLSHIVPGDPHEPEPLPEDVQ 84
 QY 58 KTSNADVIIFYNGINLEDEGQAMFTKLVNNAOKTKNKKDYFAVSDGIDVITYLEGASERKED 117
 Db 85 KTSKADLIFYNGVNLXGTGNMFTKLVXKXANKKEKNDFFASDGDIDVITYLEGOSERKED 144
 QY 118 PHAMINLENGIITYAKNIETKMLAEKDPDNKKFYKENDKYIEKIDSLDEAKSKFASIPND 177
 Db 145 PHAMINLENGIITYAKNIETKMLAEKDPDNKKFYKENDKYIEKIDSLDEAKSKFASIPND 204
 QY 178 KKLIVTSEGCERYFSKAYGVPASVAYIWEINTEEGTPDQISSLEIKLYIKPSALFEVSSV 237
 Db 205 KKLIVTSEGCERYFSKAYGVPASVAYIWEINTEEGTPDQIKHLVERLRTTKVPSLFVSSV 264
 QY 238 DRPMETVSKDSGIPYISEIFTDSIAKKGKPGDSYAMKKNLKDISEGL 287
 Db 265 DRPMETVSKDSGIPYISEIFTDSIAKKGKPGDSYAMKKNLKDISEGL 314

Search completed: May 21, 2003, 10:58:32
 Job time : 79 secs



GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 10:56:41 ; Search time 37 Seconds
(without alignments)
1033.593 Million cell updates/sec

Title: US-09-869-677A-2

Perfect score: 287
Sequence: 1 SSTGAKTAKSDKLKVAATNS.....PGDSYAMKKNLIDKISGL 287

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A.GeneSeq_101002:*
1: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	287	100.0	287	21 AAB07436	A streptococcal 11
2	231	80.5	274	23 AAB27001	Streptococcus poly
3	231	80.5	306	23 AAB66911	Streptococcus pyog
4	40	13.9	168	23 AAB27434	Streptococcus poly
5	40	13.9	308	23 AAB30114	Streptococcus poly
6	28	9.8	310	14 AAB37495	Pneumococcal fibr
7	27	9.4	289	19 AAB55072	Streptococcus pneu
8	27	9.4	289	23 AAB54566	S. pneumoniae SP01
9	27	9.4	309	20 AAY30350	37. kDa pneumococa
10	27	9.4	309	20 AAB82496	S. pneumoniae 37-k

11	27	9.4	309	22 AAB01906	Streptococcus pneu
12	27	9.4	309	23 AAB19238	S. pneumoniae pneu
13	27	9.4	310	21 AAB81668	Streptococcus pneu
14	21	7.3	22	21 AAB07437	Antigenic peptide
15	20	7.0	293	20 AAY00257	Enterococcus faeca
16	20	7.0	293	23 AAB43476	E faecalis EF132 a
17	20	7.0	316	20 AAY00256	Enterococcus faeca
18	20	7.0	316	23 AAB43475	E faecalis EF132 p
19	19	6.6	20	21 AAB07440	Antigenic peptide
20	19	6.6	313	23 AAB54653	Lactococcus lactis
21	17	5.9	18	21 AAB07439	Antigenic peptide
22	16	5.6	17	21 AAB07438	Antigenic peptide
23	15	5.2	27	20 AAY29825	37 kDa protein (1-
24	12	4.2	13	21 AAB07441	Antigenic peptide
25	11	3.8	289	20 AAY00023	Enterococcus faeca
26	11	3.8	289	23 AAB43472	E faecalis EF008 a
27	11	3.8	308	20 AAY00022	Enterococcus faeca
28	11	3.8	308	23 AAB43471	E faecalis EF008 p
29	9	3.1	310	23 AAB49144	Listeria monocyclo
30	9	2.8	54	22 AAV46618	Propionibacterium
31	8	2.8	196	23 AAB04023	Human ORF protein
32	8	2.8	255	22 AAV36295	Pseudomonas aerugi
33	8	2.8	269	22 AAB70862	Drosophila melanog
34	8	2.8	309	18 AAW26367	Staphylococcus aur
35	8	2.8	316	23 AAB40702	Staphylococcus epi
36	7	2.4	60	8 AAB71178	Bacillus exoprotaea
37	7	2.4	98	21 AAG20210	Arabidopsis thalia
38	7	2.4	98	23 AAB34948	Human ORF921 prot
39	7	2.4	109	21 AAG20209	Arabidopsis thalia
40	7	2.4	135	22 ABB52551	Escherichia coli p
41	7	2.4	142	18 AAB14554	Streptococcus pneu
42	7	2.4	166	22 AAG81390	Human AFP protein
43	7	2.4	184	22 ABB59822	Drosophila melanog
44	7	2.4	188	21 AAG11073	Arabidopsis thalia
45	7	2.4	188	21 AAG50188	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAB07436
AAB07436 standard; Protein; 287 AA.
XX
AC AAB07436;
XX
DT 20-OCT-2000 (first entry)
XX
A streptococcal lipoprotein of an ABC transporter protein.
XX
Lipoprotein; ATP-binding cassette transporter; Mtsa; immune response;
KW vaccine.
KW Streptococcus pyogenes.
OS
XX
PN WO2000040729-A1.
XX
PD 13-JUL-2000.
XX
PF 30-DEC-1999; 99WO-GH04445.
XX
PR 31-DEC-1998; 98GB-0028880.
XX
PA (ACT1-) ACTINOVA LTD.
XX
PI Bjoerck L, Janulczyk R;
XX
DR WPI: 2000-465989/40.
XX
DR N-PSDB; AAB58808.
XX
PT New polypeptide comprising an ATP-binding cassette transporter of
PT Streptococcus pyogenes for use as a vaccine against Streptococcus and
PT for assays that detect immune reactivity to the polypeptides in animals

PT and humans -
 XX Claim 1; Page 47-48; 55pp; English.
 XX
 CC The present sequence represents a Streptococcus pyogenes polypeptide
 CC which is lipoprotein of an ATP-binding cassette transporter (Msa).
 CC The polypeptide generates a protective immune response to
 CC Streptococcus, preferably group A, such as S. pyogenes. It
 CC is used to prepare a vaccine against Streptococcus. The new
 CC polypeptides may be used in serological or cell mediated immune assays
 CC for the detection of immune reactivity to the polypeptides in animals
 CC and humans. Antibodies to Msa protein can be detected using an
 CC immunosassay with the polypeptides.
 XX
 SQ Sequence 287 AA;
 Query Match 100.0%; Score 287; DB 21; Length 287;
 Best Local Similarity 100.0%; Pred. No. 5.4e-281;
 Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSTGAKTAKSDKLKVAATNSIIADMTKAIAGDKIDLSIIVPGDPHREPLPEDAEKTS 60
 DB 1 SSTGAKTAKSDKLKVAATNSIIADMTKAIAGDKIDLSIIVPGDPHREPLPEDAEKTS 60
 QY 61 NADYIFNGINLEDEGGAMFTKLVNAOKTKNDYFAVSDGIDVYILEGASEKGEDPHA 120
 DB 61 NADYIFNGINLEDEGGAMFTKLVNAOKTKNDYFAVSDGIDVYILEGASEKGEDPHA 120
 QY 121 WLNLENGIYSKNIQAOLIAKDPKNETEYKNAVAKLEKLDKREKSFDAIAENKKL 180
 DB 121 WLNLENGIYSKNIQAOLIAKDPKNETEYKNAVAKLEKLDKREKSFDAIAENKKL 180
 QY 181 IVTSEGGFKYFSKAYGVPASAYIWEINTEEGTPOQISSLEKLVKIPSALEFVSSVDR 240
 DB 181 IVTSEGGFKYFSKAYGVPASAYIWEINTEEGTPOQISSLEKLVKIPSALEFVSSVDR 240
 QY 241 PMETVSKDSGIPYSEIFTSIAKKGKPGDSYAMMKWNLDKISEGL 287
 DB 241 PMETVSKDSGIPYSEIFTSIAKKGKPGDSYAMMKWNLDKISEGL 287
 RESULT 2
 ABP27001 standard; Protein: 274 AA.
 XX
 AC ABP27001;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 3178.
 XX
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 PN WO200234771-A2.
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001MO-GH04789.
 XX
 PR 27-OCT-2000; 2000GB-0026333.
 XX
 PR 24-NOV-2000; 2000GB-0028727.
 XX
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 XX (GENO-) INST GENOMIC RES.
 XX Telford J, Maignani V, Margalit Ros YI, Grandi G, Fraser C;
 PI Tetelin H;
 XX

DR WPI: 2002-352536/38.
 DR N-PSDB; ABN67632.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 PS Claim 1; Page 3471; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC Streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN65044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX
 SQ Sequence 274 AA;
 Query Match 80.5%; Score 231; DB 23; Length 274;
 Best Local Similarity 100.0%; Pred. No. 1.6e-224;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 57 EKTSNADYIFNGINLEDEGGAMFTKLVNAOKTKNDYFAVSDGIDVYILEGASEKGE 116
 DB 42 EKTSNADYIFNGINLEDEGGAMFTKLVNAOKTKNDYFAVSDGIDVYILEGASEKGE 101
 QY 117 DPHAMLENGIYSKNIQAOLIAKDPKNETEYKNAVAKLEKLDKREKSFDAIAE 176
 DB 102 DPHAMLENGIYSKNIQAOLIAKDPKNETEYKNAVAKLEKLDKREKSFDAIAE 161
 QY 177 NKLIIVTSEGGFKYFSKAYGVPASAYIWEINTEEGTPOQISSLEKLVKIPSALEFVSS 236
 DB 162 NKLIIVTSEGGFKYFSKAYGVPASAYIWEINTEEGTPOQISSLEKLVKIPSALEFVSS 221
 QY 237 VDRPMETVSKDSGIPYSEIFTSIAKKGKPGDSYAMMKWNLDKISEGL 287
 DB 222 VDRPMETVSKDSGIPYSEIFTSIAKKGKPGDSYAMMKWNLDKISEGL 272
 RESULT 3
 ABG66911 standard; Protein: 306 AA.
 XX
 ID ABG66911
 XX
 AC ABG66911;
 XX
 DT 24-SEP-2002 (first entry)
 XX
 DE Streptococcus pyogenes strain M1 BVH-p3.
 XX
 XX Streptococcus polypeptide; gene therapy; pharyngitis; erysipelas;
 KW impetigo; scarlet fever; invasive disease; bacteraemia;
 KM necrotising fasciitis; toxic shock; bacterial infection; vaccine.
 XX
 OS Streptococcus pyogenes.
 XX
 PN WO200250107-A2.
 PD 27-JUN-2002.
 XX
 PF 21-DEC-2001; 2001MO-CA01853.
 XX
 PR 21-DEC-2000; 2000US-256940P.
 XX

XX (SHIR-) SHIRE BIOCHEM INC.
PA Martin D, Brodeur BR, Hamel J, Rioux S, Rheault P;
XX WPI: 2002-500623/53.
DR N-PSDB: ABK95351.
XX Novel polypeptide useful for diagnostic Streptococcal bacterial
PT infection in a host susceptible to Streptococcal infection -
PS
XX Claim 1; Fig 4; 55pp; English.
XX The invention describes an isolated Streptococcal polypeptide (I).
CC (I) is useful for diagnosing streptococcal bacterial infection in a host
CC susceptible to streptococcal infection by obtaining a biological sample
CC from a host, incubating an antibody or its fragment reactive with (I)
CC with the biological sample to form a mixture, and detecting specifically
CC bound antibody or antigen or bound fragment in the mixture which
CC indicates the presence of Streptococci or an antibody specific to
CC streptococci. A composition comprising (I) is useful for treatment (e.g.
CC by gene therapy) of pharyngitis, erysipelas and impetigo, scarlet fever,
CC and invasive diseases such as bacteraemia, necrotising fasciitis and
CC toxic shock. The composition is also useful for therapeutic or
CC prophylactic treatment of Streptococcus pyogenes bacterial infection in a
CC host susceptible to Streptococcus pyogenes infection. This is the amino
CC acid sequence of Streptococcus pyogenes BVH antigen used as a vaccine
CC component for therapy and/or prophylaxis.
XX
SQ Sequence 306 AA:
Query Match 80.5%; Score 231; DB 23; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.8e-224; Indels 0; Gaps 0;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 57 EKSNADVIFYNGINLEDGQAFETKLVKNAQKTKNKNDYFVSDGIDVITYLEGASKEGKE 116
DB 74 EKSNADVIFYNGINLEDGQAFETKLVKNAQKTKNKNDYFVSDGIDVITYLEGASKEGKE 133
OY 117 DPHAMWLENGIITYSKNIAKQIAKDPKNEYEKNIKAVAKLEKIDRAKSPDAIAE 176
DB 134 DPHAMWLENGIITYSKNIAKQIAKDPKNEYEKNIKAVAKLEKIDRAKSPDAIAE 193
OY 177 NKRLIVTSEGCFFYFSKAYGVPASAYIWEINTEEGTPDQISLIEKLVKPSALFEVSS 236
DB 194 NKRLIVTSEGCFFYFSKAYGVPASAYIWEINTEEGTPDQISLIEKLVKPSALFEVSS 253
OY 237 VDRPMETVSKDSGIPYSEIFETDSIAKGRPGDSYVAMKWNLDKISEGI 287
DB 254 VDRPMETVSKDSGIPYSEIFETDSIAKGRPGDSYVAMKWNLDKISEGI 304
RESULT 4
ABP27434 ID ABP27434 standard; Protein: 168 AA.
XX
AC ABP27434:
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 4044.
XX
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PR 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tectelin H;
XX WPI: 2002-352536/38.
DR N-PSDB: ABN68065.
XX
XX New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
PS
XX Claim 1; Page 3560; 4525pp; English.
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 168 AA:
Query Match 13.9%; Score 40; DB 23; Length 168;
Best Local Similarity 100.0%; Pred. No. 5.1e-32; Indels 0; Gaps 0;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 177 NKRLIVTSEGCFFYFSKAYGVPASAYIWEINTEEGTPDQI 216
DB 56 NKRLIVTSEGCFFYFSKAYGVPASAYIWEINTEEGTPDQI 95
RESULT 5
ABP30114 ID ABP30114 standard; Protein: 308 AA.
XX
AC ABP30114:
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 9404.
XX
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PR 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.

07-MAR-2001; 2001GB-0005640.
 (CHIR-) CHIRON SPA.
 (GENO-) INST-GENOMIC RES.
 Telford J, Maslgnani V, Margarit Ros YI, Grandi G, Fraser C;
 Tettelin H;
 WPI: 2002-352536/38.
 N-PSDB: ABR70745.
 New Streptococcus protein for the treatment or prevention of infection or disease caused by Streptococcus bacteria, such as meningitis, and for detecting a compound that binds to the protein -
 Claim 1; Page 4066; 4525pp; English.
 The invention relates to a protein (ABP25413-ABP30895) from group B Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in the specification. The proteins have antibacterial and anti-inflammatory activity. (1), nucleic acids encoding (1), ABR66044-ABN71526 and antibodies that bind (1) are used in the manufacture of medicaments for the treatment or prevention of infection or disease caused by Streptococcus bacteria, particularly S. agalactiae and S. pyogenes. Nucleic acids encoding (1) is used to determine whether a compound binds to (1). A composition comprising (1) or a nucleic acid encoding (1), may be used as a vaccine or diagnostic composition. The disease caused by Streptococcus that is prevented or treated may be meningitis. Nucleic acid encoding (1) may be used to recombinantly produce (1) and may be used in gene therapy. Antibodies to (1) are used for affinity chromatography, immunoassays, and distinguishing/identifying Streptococcus proteins.
 Sequence 308 AA;
 Query Match 13.9%; Score 40; DB 23; Length 308;
 Best Local Similarity 100.0%; Pred. No. 8, 8e-32;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 177 NKKLIVTSEGCFKFKYKAGVPSAYIWEINTEEGTPOI 216
 196 NKKLIVTSEGCFKFKYKAGVPSAYIWEINTEEGTPOI 235
 RESULT 6
 AAR37495
 ID AAR37495 standard; Protein; 310 AA.
 AAR37495;
 17-SEP-1993 (first entry)
 Pneumococcal fimbrial protein A.
 Pneumonia; vaccine; combination; B-cell; T-cell; stimulation;
 diagnosis; pneumococcal disease; children; adults; Pfpa.
 Streptococcus pneumoniae.
 WO9310238-A.
 27-MAY-1993.
 16-NOV-1992; 92WO-US09522.
 14-NOV-1991; 91US-0791377.
 03-JAN-1992; 92US-0816286.
 (USSH) US DEPT HEALTH & HUMAN SERVICE.
 O'Connor SP, Russell H, Sampson J, Tharpe JA;

1993-182553/22.
 N-PSDB: AAQ42845.
 Vaccine to prevent pneumococcal pneumonia - comprises pneumococcal fimbrial protein A as conjugate with carrier e.g. keyhole limpet haemocyanin
 Claim 27; Page 33-34; 50pp; English.
 The sequence is that of pneumococcal fimbrial protein A (Pfpa) which can be produced using recombinant technology. It can be used in a vaccine which is capable of eliciting a combination of B- and T-cell stimulation, and for diagnosis of pneumococcal disease in children and adults.
 Sequence 310 AA;
 Query Match 9.8%; Score 28; DB 14; Length 310;
 Best Local Similarity 100.0%; Pred. No. 1, 1e-19;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 181 IVTSEGCFKFKYKAGVPSAYIWEINTE 208
 202 IVTSEGCFKFKYKAGVPSAYIWEINTE 229
 RESULT 7
 AAM55072
 ID AAM55072 standard; Protein; 289 AA.
 AAM55072;
 02-OCT-1998 (first entry)
 Streptococcus pneumoniae sp0013 protein.
 Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
 detection; pneumonia; otitis media; meningitis.
 Streptococcus pneumoniae.
 WO9818930-A2.
 07-MAY-1998.
 30-OCT-1997; 97WO-US19422.
 31-OCT-1996; 96US-0029960.
 (HUMA-) HUMAN GENOME SCI INC.
 Choi GH, Hromocky A, Johnson LS, Kunsch CA;
 WPI: 1998-272224/24.
 N-PSDB: AAV27333.
 Nucleic acid encoding antigenic peptide(s) from Streptococcus pneumoniae - or their epitope-containing fragments, useful in protective or therapeutic vaccines, and for diagnosis
 Claim 11; Page 52; 118pp; English.
 The present sequence represents a protein from Streptococcus pneumoniae. The nucleic acid sequence encoding the Streptococcus pneumoniae protein can be useful in vaccines for inducing protective antibodies against Streptococcus pneumoniae, for treatment or prevention of infection e.g. pneumonia, otitis media or meningitis. Probes based on the nucleic acid are used to detect Streptococcus infection (by usual hybridisation or amplification methods), also for isolating Streptococcus genes or their allelic variants. The protein can be used similarly to detect specific antibodies in standard immunoassays, especially for diagnosing or monitoring infections. Antibodies which bind the protein are used to

KW diagnostic; immunoassay; treatment; infection; anti-idiotype.
 XX Streptococcus pneumoniae.
 XX US5854416-A.
 PN 29-DEC-1998.
 PD 17-SEP-1996; 96US-0715131.
 XX 17-SEP-1996; 96US-0715131.
 PR 17-SEP-1996; 96US-0715131.
 PR 04-APR-1994; 94US-0222179.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Ades EW, Carlone GM, Russell H, Sampson JS, Tharpe JA;
 PI WPI; 1999-095007/08.
 DR N-PSDB; AAV73914.
 XX Nucleic acid encoding the 37 kDa, surface adhesion A of Streptococcus
 PT pneumoniae - useful diagnostically and for production of
 PT recombinant polypeptides
 PS Claim 1: Column 33-34; 20pp; English.
 XX This sequence represents a Streptococcus pneumoniae 37-kDa surface
 CC adhesion A protein. This encoding nucleic acid can be used in methods to
 CC express recombinant protein, as a source of primers for amplification (to
 CC identify and isolate related sequences), e.g. allelic variants) or probes
 CC for nucleic acid hybridization tests for detecting S. pneumoniae, and in
 CC DNA vaccines. This protein and its fragments can be used to raise
 CC antibodies, in vaccines and for detecting S. pneumoniae (by reaction with
 CC specific antibodies). Antibodies are useful in diagnostic immunoassays,
 CC to treat infections and to raise anti-idiotype antibodies for use in
 CC vaccines. This protein is very highly conserved between the different
 CC serotypes of S. pneumoniae so is an excellent candidate for vaccine
 CC development.
 SO Sequence 309 AA;
 Query Match 9.4%; Score 27; DB 20; Length 309;
 Best Local Similarity 100.0%; Pred. No. 1.2e-18;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 29 IAGDKIDLHSTIVPIGDPHNEYEPLEPED 55
 DB 49 IAGDKIDLHSTIVPIGDPHNEYEPLEPED 75
 RESULT 11
 AAU01906
 ID AAU01906 standard; protein; 309 AA.
 AC AAU01906;
 XX 29-AUG-2001 (first entry)
 DT 29-AUG-2001 (first entry)
 XX Streptococcus pneumoniae 37kDa surface adhesion A protein.
 DE Streptococcus pneumoniae 37kDa surface adhesion A protein.
 KW 37-kDa surface adhesion A; pneumococcal disease; vaccine; treatment;
 KM infection.
 XX Streptococcus pneumoniae.
 OS US6212884-A1.
 PN 17-APR-2001.
 PD 28-DEC-1998; 98US-0221753.
 PF 17-SEP-1996; 96US-0715131.
 PR 17-SEP-1996; 96US-0715131.

PR 17-SEP-1991; 91US-0791377.
 PR 04-APR-1994; 94US-0222179.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Sampson JS, Russell H, Tharpe JA, Adep EW, Carlone GM;
 PI WPI; 2001-289821/30.
 DR N-PSDB; AAS03810.
 XX New 37-kDa pneumococcal surface adhesion A protein from Streptococcus
 PT pneumoniae, useful as a vaccine for treating or preventing infections
 PT caused by Streptococcus pneumoniae -
 PS Disclosure; Column 33-34; 20pp; English.
 XX The sequence represents Streptococcus pneumoniae 37-kDa surface adhesion A
 CC protein. Infection by S. pneumoniae leads to pneumococcal disease. The
 CC 37-kDa surface adhesion A protein and its corresponding DNA can be used as
 CC a vaccine component for treatment and prevention of pneumococcal disease,
 CC as well as a reagent for identifying host antibodies raised against
 CC S. pneumoniae during infection. The protein may also be used to detect
 CC the presence of S. pneumoniae. The nucleic acids can be used as primers
 CC for amplifying nucleic acids from other strains of S. pneumoniae to
 CC isolate allelic variants of the protein, or for reverse transcription
 CC techniques, and as probes for use in detection techniques such as nucleic
 CC acid hybridization.
 SO Sequence 309 AA;
 Query Match 9.4%; Score 27; DB 22; Length 309;
 Best Local Similarity 100.0%; Pred. No. 1.2e-18;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 29 IAGDKIDLHSTIVPIGDPHNEYEPLEPED 55
 DB 49 IAGDKIDLHSTIVPIGDPHNEYEPLEPED 75
 RESULT 12
 AAE19238
 ID AAE19238 standard; protein; 309 AA.
 AC AAE19238;
 XX 21-MAY-2002 (first entry)
 DT 21-MAY-2002 (first entry)
 XX S. pneumoniae pneumococcal surface adhesion protein A (psaa).
 DE Multiple antigenic peptide; MAP; immunogenic; immunity; infection;
 KW pneumococcal surface adhesion protein A; psaa; antibacterial.
 OS Streptococcus pneumoniae.
 PN WO200204497-A2.
 PD 17-JAN-2002.
 XX 10-JUL-2001; 2001WO-US21626.
 PF 10-JUL-2000; 2000US-0613092.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Ades EW, Johnson SE, Jue DL, Sampson JS, Carlone GM;
 PI WPI; 2002-195762/25.
 DR N-PSDB; AAD30661.
 XX New multiple antigenic peptide for immunizing against streptococcal
 PT infections, binds to monoclonal antibody obtained in response to
 PT immunizing an animal with pneumococcal surface adhesion protein A or
 PT its fragment

XX PS Disclosure; Page 84; 86pp; English.
XX CC The invention relates to multiple antigenic peptides (MAP) immunogenic
CC against Streptococcus pneumoniae. MAP binds to monoclonal antibody
CC obtained in response to immunising an animal with pneumococcal surface
CC adhesion protein A (PsaA) or its fragment. MAP is useful for conferring
CC protective immunity against S. pneumoniae infection in a subject. The
CC present sequence is a 37-kDa Streptococcus pneumoniae PsaA protein.
XX SQ Sequence 309 AA;
Query Match 9.4%; Score 27; DB 23; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.2e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 29 IAGDKIDLHSIVPIGDPHEPEPLPED 55
DB 49 IAGDKIDLHSIVPIGDPHEPEPLPED 75
RESULT 13
AAV81668
ID AAV81668 standard; Protein; 310 AA.
XX AC AAV81668;
XX DT 24-MAY-2000 (first entry)
XX DE Streptococcus pneumoniae protein sequence ID205 - 4118.1.
XX KM Streptococcus pneumoniae; vaccine; screening; protein antigen;
XX KM antibacterial; antiinflammatory; meningitis; infection; diagnosis;
XX KM pneumococcal disease.
XX OS Streptococcus pneumoniae.
XX PN WO200006737-A2.
XX PD 10-FEB-2000.
XX PF 27-JUL-1999; 99WO-GB02451.
XX PR 27-JUL-1998; 98GB-0016337.
XX PR 19-MAR-1999; 99US-0125164.
XX PA (MICR-) MICROBIAL TECHNIQS LTD.
XX PI Gilbert CFG, Hansdro PM;
XX DR WPI; 2000-195300/17.
XX PT New Streptococcal protein, useful as a vaccine, for diagnosis of
XX PT pneumococcal diseases and for screening agents capable of antagonizing
XX PT or inhibiting expression of the protein
XX PS Claim 6; Page 90; 108pp; English.
XX AA AAV81501 to AAV81679 represent specifically claimed protein sequences
CC isolated from Streptococcus pneumoniae. AA05407 to AA05590 represent
CC specifically claimed nucleotide sequences isolated from S. pneumoniae.
CC The sequences have antibacterial and antiinflammatory properties.
CC The protein sequences, and fragments of them, are useful as immunogens
CC and/or adjuvants. The nucleotide sequences can be used in vaccines and in
CC diagnostic assays. The proteins and nucleotides can be useful for the
CC detection and diagnosis of S. pneumoniae. The protein sequences are also
CC useful for screening an agent capable of antagonising, inhibiting or
CC interfering with the function or expression of the proteins in which the
CC agent is useful for treatment or prophylaxis of S. pneumoniae infection
CC and meningitis. AA05591 to AA05614 represent primers used in the
CC exemplification of the present invention.
XX SQ Sequence 310 AA;

Query Match 9.4%; Score 27; DB 21; Length 310;
Best Local Similarity 100.0%; Pred. No. 1.2e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 29 IAGDKIDLHSIVPIGDPHEPEPLPED 55
DB 49 IAGDKIDLHSIVPIGDPHEPEPLPED 75
RESULT 14
AAB07437
ID AAB07437 standard; Peptide; 22 AA.
XX AC AAB07437;
XX DT 20-OCT-2000 (first entry)
XX DE Antigenic peptide from a lipoprotein of an ABC transporter protein.
XX KM Lipoprotein; ATP-binding cassette transporter; MtsA; immune response;
XX KM vaccine.
XX OS Streptococcus pyogenes.
XX FH Key location/qualifiers
XX FT Misc-difference /note- "this residue is not part of the native sequence"
XX PN WO2000040729-A1.
XX PD 13-JUL-2000.
XX PF 30-DEC-1999; 99WO-GB04445.
XX PR 31-DEC-1998; 98GB-0028880.
XX PA (ACT-) ACTINOVA LTD.
XX PI Bjoerck L, Janulczyk R;
XX DR WPI; 2000-465989/40.
XX PT New polypeptide comprising an ATP-binding cassette transporter of
XX PT Streptococcus pyogenes for use as a vaccine against Streptococcus and
XX PT for assays that detect immune reactivity to the polypeptides in animals
XX PT and humans
XX PS Example 7; Page 31; 55pp; English.
XX CC AAB07437-41 represent antigenic peptides derived from a Streptococcus
CC pyogenes polypeptide which is lipoprotein of an ATP-binding cassette
CC transporter (MtsA). The polypeptide generates a protective immune
CC response to Streptococcus, preferably group A, such as S. pyogenes.
CC It is used to prepare a vaccine against Streptococcus. The new
CC polypeptides may be used in serological or cell mediated immune
CC assays for the detection of immune reactivity to the polypeptides in
CC animals and humans. Antibodies to MtsA protein can be detected using
CC an immunoassay with the polypeptides.
XX SQ Sequence 22 AA;
Query Match 7.3%; Score 21; DB 21; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 259 TDSIAKKGKPGDSYAMMKWN 279
DB 2 TDSIAKKGKPGDSYAMMKWN 22
RESULT 15
AAV00257

ID AAY00257 standard; Protein; 293 AA.
 XX AAY00257;
 AC
 XX 20-APR-1999 (first entry)
 DT
 XX Enterococcus faecalis antigenic polypeptide fragment EF132.
 DE
 XX Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
 KW detection; attenuation; antigenic.
 RM Enterococcus faecalis.
 XX
 OS WO9850554-A2.
 PN
 XX 12-NOV-1998.
 PD
 XX 04-MAY-1998; 98WO-US08959.
 PE
 XX 14-NOV-1997; 97US-0066009.
 PR 06-MAY-1997; 97US-0044031.
 PR 16-MAY-1997; 97US-0046655.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Bailey C, Choi GH, Hromocky J A, Kunsch CA;
 XI WPI: 1999-070095/06.
 DR N-PSDB; AAX20247.
 DR
 XX
 PT New isolated Enterococcus faecalis polynucleotides - used to develop
 PT products for the detection of Enterococcus and for use in vaccines
 PT for prevention or attenuation of Enterococcus infection
 XX
 PS Claim 9; Page 243; 301pp; English.
 PS
 CC The present sequence represents an antigenic polypeptide fragment
 CC isolated from Enterococcus faecalis. The present invention describes
 CC genes, proteins and antigenic polypeptides isolated from E. faecalis.
 CC The proteins can be used in vaccines for preventing or attenuating an
 CC infection caused by a member of the Enterococcus genus in an animal.
 CC They can also be used for detecting Enterococcus antibodies in a sample.
 CC The nucleotide sequences can be used for detecting Enterococcus nucleic
 CC acids. Products from the present invention can also be used for
 CC screening compounds to identify agonists and antagonists of E. faecalis
 CC protein activity.
 CC
 SQ Sequence 293 AA;
 SQ
 Query Match 7.0%; Score 20; DB 20; Length 293;
 Best Local Similarity 100.0%; Pred. No. 1.3e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 111 SEKGKEDPHAMLNENGIIT 130
 DB 115 SEKGKEDPHAMLNENGIIT 134
 DE
 XX
 AC AAY00256 standard; Protein; 293 AA.
 ID AAY00256;
 AC
 XX 05-AUG-2002 (first entry)
 DT
 XX E faecalis EF132 antigenic fragment.
 DE
 XX Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.
 KW Enterococcus faecalis.
 OS US2002045737-A1.
 PN

XX 18-APR-2002.
 PD
 XX 04-MAY-1998; 98US-0071035.
 PE
 XX 04-MAY-1998; 98US-0071035.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 PI Choi GH, Bailey C, Hromocky J A, Kunsch CA;
 XI WPI: 2002-425450/45.
 DR N-PSDB; ABN98232.
 DR
 XX
 PT New genes and polypeptides from Enterococcus faecalis, useful as
 PT vaccines for preventing, treating or attenuating an infection caused by
 PT a member of the Enterococcus genus in an animal, particularly E.
 PT faecalis -
 PS
 PS Claim 9; Page 239; 255pp; English.
 PS
 CC The present invention provides the protein and coding sequences of a
 CC number of polypeptides from Enterococcus faecalis. The proteins can be
 CC used as vaccines for preventing or attenuating an infection caused by a
 CC member of the Enterococcus genus in an animal, particularly E. faecalis.
 CC The polynucleotide is also useful for preventing or treating E. faecalis
 CC infection. The present sequence is a protein of the invention.
 CC
 SQ Sequence 293 AA;
 SQ
 Query Match 7.0%; Score 20; DB 23; Length 293;
 Best Local Similarity 100.0%; Pred. No. 1.3e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 111 SEKGKEDPHAMLNENGIIT 130
 DB 115 SEKGKEDPHAMLNENGIIT 134
 DE
 XX
 AC AAY00256 standard; Protein; 316 AA.
 ID AAY00256;
 AC
 XX 20-APR-1999 (first entry)
 DT
 XX Enterococcus faecalis protein EF132.
 DE
 XX Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
 KW detection; attenuation; antigenic.
 RM Enterococcus faecalis.
 XX
 OS WO9850554-A2.
 PN
 XX 12-NOV-1998.
 PD
 XX 04-MAY-1998; 98WO-US08959.
 PE
 XX 14-NOV-1997; 97US-0066009.
 PR 06-MAY-1997; 97US-0044031.
 PR 16-MAY-1997; 97US-0046655.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Bailey C, Choi GH, Hromocky J A, Kunsch CA;
 XI WPI: 1999-070095/06.
 DR N-PSDB; AAX20246.
 DR
 XX
 PT New isolated Enterococcus faecalis polynucleotides - used to develop
 PT products for the detection of Enterococcus and for use in vaccines


PT for prevention or attenuation of Enterococcus infection
 XX
 PS Claim 9; Page 243; 301pp; English.
 CC The present sequence represents a protein isolated from
 CC Enterococcus faecalis. The present invention describes genes, proteins
 CC and antigenic polypeptides isolated from E. faecalis. The proteins can
 CC be used in vaccines for preventing or attenuating an infection caused
 CC by a member of the Enterococcus genus in an animal. They can also be
 CC used for detecting Enterococcus antibodies in a sample. The nucleotide
 CC sequences can be used for detecting Enterococcus nucleic acids.
 CC Products from the present invention can also be used for screening
 CC compounds to identify agonists and antagonists of E. faecalis protein
 CC activity.
 CC
 SQ Sequence 316 AA;
 7.0%; Score 20; DB 20; Length 316;
 Best Local Similarity 100.0%; Pred. No. 1.4e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 111 SEKGKEDPHAWLNLENGIT 130
 DB 138 SEKGKEDPHAWLNLENGIT 157
 RESULT 18
 ABP3475
 ID ABP3475 standard; Protein; 316 AA.
 XX
 AC ABP3475;
 XX
 DT 05-AUG-2002 (first entry)
 XX
 DE E faecalis EF132 protein.
 XX
 KM Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.
 XX
 OS Enterococcus faecalis.
 XX
 PN US2002045737-A1.
 PD 18-APR-2002.
 XX
 PE 04-MAY-1998; 9805-0071035.
 XX
 PR 04-MAY-1998; 9805-0071035.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Choi GH, Bailey C, Hromocky J A, Kunsch CA;
 XX
 DR WPI; 2002-425450/45.
 DR N-PSDB; ABN98231.
 XX
 PT New genes and polypeptides from Enterococcus faecalis, useful as
 PT vaccines for preventing, treating or attenuating an infection caused by
 PT a member of the Enterococcus genus in an animal, particularly E.
 PT faecalis -
 PT
 PS Claim 9; Page 238; 255pp; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of polypeptides from Enterococcus faecalis. The proteins can be
 CC used as vaccines for preventing or attenuating an infection caused by a
 CC member of the Enterococcus genus in an animal, particularly E. faecalis.
 CC The polynucleotide is also useful for preventing or treating E. faecalis
 CC infection. The present sequence is a protein of the invention.
 CC
 SQ Sequence 316 AA;
 7.0%; Score 20; DB 23; Length 316;
 Best Local Similarity 100.0%; Pred. No. 1.4e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 111 SEKGKEDPHAWLNLENGIT 130
 DB 138 SEKGKEDPHAWLNLENGIT 157
 RESULT 19
 AAB07440
 ID AAB07440 standard; Peptide; 20 AA.
 XX
 AC AAB07440;
 XX
 DT 20-OCT-2000 (first entry)
 XX
 DE Antigenic peptide from a lipoprotein of an ABC transporter protein.
 XX
 KW Lipoprotein; ATP-binding cassette transporter; MtsA; immune response;
 XX vaccine.
 XX
 OS Streptococcus pyogenes.
 XX
 FH Key Location/Qualifiers
 FT MISC-difference 1
 FT note="this residue is not part of the native sequence"
 XX
 PN W0200040729-A1.
 PD 13-JUL-2000.
 XX
 PE 30-DEC-1999; 99WO-GB04445.
 XX
 PR 31-DEC-1998; 98GB-0028880.
 XX
 PA (ACTI-) ACTINOVA LTD.
 XX
 PI Bjoerck L, Janulczyk R;
 XX
 DR WPI; 2000-465989/40.
 XX
 PT New polypeptide comprising an ATP-binding cassette transporter of
 PT Streptococcus pyogenes for use as a vaccine against Streptococcus and
 PT for assays that detect immune reactivity to the polypeptides in animals
 PT and humans -
 PT
 PS Example 7; Page 31; 55pp; English.
 XX
 CC AAB07437-41 represent antigenic peptides derived from a Streptococcus
 CC pyogenes polypeptide which is lipoprotein of an ATP-binding cassette
 CC transporter (MtsA). The polypeptide generates a protective immune
 CC response to Streptococcus, preferably group A, such as S. pyogenes.
 CC It is used to prepare a vaccine against Streptococcus. The new
 CC polypeptides may be used in serological or cell mediated immune
 CC assays for the detection of immune reactivity to the polypeptides in
 CC animals and humans. Antibodies to MtsA protein can be detected using
 CC an immunoassay with the polypeptides.
 CC
 SQ Sequence 20 AA;
 6.6%; Score 19; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.1e-11;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 204 EINTREESTPQISSLIK 222
 DB 2 EINTREESTPQISSLIK 20
 RESULT 20
 ABB54653
 ID ABB54653 standard; Protein; 313 AA.
 AC ABB54653;
 7.0%; Score 20; DB 23; Length 316;
 Best Local Similarity 100.0%; Pred. No. 1.4e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 16-MAY-2002 (first entry)
 DT Lactococcus lactis protein mtsa.
 XX DE Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX KM Lactococcus lactis IL1403.
 XX OS
 XX ER2807448-A1.
 XX PN 12-OCT-2001.
 XX PD
 XX 11-APR-2000; 2000FR-0004630.
 XX PF
 XX 11-APR-2000; 2000FR-0004630.
 XX PR
 XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX PA
 XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 XX PI
 XX WPI: 2002-043418/06.
 XX DR
 XX
 XX New nucleotide sequence useful in the identification or Lactococcus
 PT lactis and related species -
 PT
 XX Claim 6; SEQ ID No 1355; 2504bp; French.
 PS
 XX The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (AB53300-AB53621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO200177334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 313 AA;
 Query Match 6.6%; Score 19; DB 23; Length 313;
 Best Local Similarity 100.0%; Pred. No. 14e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 198 PSAYIWEINTEEGTPOI 216
 II
 DB 222 PSAYIWEINTEEGTPOI 240
 II
 RESULT 21
 AAB07439
 ID AAB07439 standard; Peptide: 18 AA.
 AC AAB07439;
 XX
 XX 20-OCT-2000 (first entry)
 DT
 XX Antigenic peptide from a lipoprotein of an ABC transporter protein.
 DE
 XX Lipoprotein; ATP-binding cassette transporter; Mtsa; Immune response;
 KW vaccine.
 KW
 XX Streptococcus pyogenes.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 1
 FT /note="this residue is not part of the native sequence"
 FT
 XX WO200040729-A1.
 PN 13-JUL-2000.
 PD

XX 30-DEC-1999; 99WO-GB04445.
 PF
 XX 31-DEC-1998; 98GB-0028880.
 XX PR
 XX (ACTI-) ACTINOVA LTD.
 XX PA
 XX Bjoerck L, Janulczyk R;
 XX PI
 XX WPI: 2000-465989/40.
 XX DR
 XX
 XX New polypeptide comprising an ATP-binding cassette transporter of
 PT Streptococcus pyogenes for use as a vaccine against Streptococcus and
 PT for assays that detect immune reactivity to the polypeptides in animals
 PT and humans -
 XX
 XX Example 7; Page 31; 55pp; English.
 PS
 XX AAB07437-41 represent antigenic peptides derived from a Streptococcus
 CC pyogenes polypeptide which is lipoprotein of an ATP-binding cassette
 CC transporter (Mtsa). The polypeptide generates a protective immune
 CC response to Streptococcus, preferably group A, such as S. pyogenes.
 CC It is used to prepare a vaccine against Streptococcus. The new
 CC polypeptides may be used in serological or cell mediated immune
 CC assays for the detection of immune reactivity to the polypeptides in
 CC animals and humans. Antibodies to Mtsa protein can be detected using
 CC an immunoassay with the polypeptides.
 CC
 SQ Sequence 18 AA;
 Query Match 5.9%; Score 17; DB 21; Length 18;
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 136 KQIADPKNKETERN 152
 II
 DB 2 KQIADPKNKETERN 18
 II

RESULT 22
 AAB07438
 ID AAB07438 standard; Peptide: 17 AA.
 AC AAB07438;
 XX
 XX 20-OCT-2000 (first entry)
 DT
 XX Antigenic peptide from a lipoprotein of an ABC transporter protein.
 DE
 XX Lipoprotein; ATP-binding cassette transporter; Mtsa; Immune response;
 KW vaccine.
 KW
 XX Streptococcus pyogenes.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 1
 FT /note="this residue is not part of the native sequence"
 FT
 XX WO200040729-A1.
 PN 13-JUL-2000.
 PD
 XX 30-DEC-1999; 99WO-GB04445.
 XX PF
 XX 31-DEC-1998; 98GB-0028880.
 XX PR
 XX (ACTI-) ACTINOVA LTD.
 XX PA
 XX Bjoerck L, Janulczyk R;
 XX PI
 XX WPI: 2000-465989/40.
 XX DR
 XX
 XX New polypeptide comprising an ATP-binding cassette transporter of

PT Streptococcus pyogenes for use as a vaccine against Streptococcus and
PT for assays that detect immune reactivity to the polypeptides in animals
PT and humans -
XX
PS Example 7; Page 31; 55pp; English.
XX
CC AAB07437-41 represent antigenic peptides derived from a Streptococcus
CC pyogenes polypeptide which is lipoprotein of an ATP-binding cassette
CC transporter (MtsA). The polypeptide generates a protective immune
CC response to Streptococcus, preferably group A, such as S. pyogenes.
CC It is used to prepare a vaccine against Streptococcus. The new
CC polypeptides may be used in serological or cell mediated immune
CC assays for the detection of immune reactivity to the polypeptides in
CC animals and humans. Antibodies to MtsA protein can be detected using
CC an immunoassay with the polypeptides.
XX
SQ Sequence 17 AA:
Query Match 5.6%; Score 16; DB 21; Length 17;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 234 ESSYDRRPMETVSKDS 249
DB 2 ESSYDRRPMETVSKDS 17
|||||
RESULT 23
AAV29825
ID AAV29825 standard; peptide; 27 AA.
XX
AC AAV29825;
XX
DT 15-NOV-1999 (first entry)
XX
DE 37 kDa protein (1-30) SP37.
XX
KW T-cell epitope; meningococcal; pneumococcal; bactericidal;
KW multi-oligosaccharide glycoconjugate bacterial meningitis vaccine;
KW pneumolysin; immunogenic; immune response; vaccination; meningitis;
KW tumour-related antigen; diagnosis; detection; anti-cancer.
XX
OS Synthetic.
XX
PN WO9942130-A1.
XX
PD 26-AUG-1999.
XX
PF 23-FEB-1999; 99WO-CA00157.
XX
PR 23-FEB-1998; 98US-0027956.
XX
PA (CONN-) CONNAUGHT LAB LTD.
XX
PI Chong P, Klein MH, Lindberg A;
XX
DR WPI; 1999-540273/45.
XX
PT Multivalent immunogenic molecule comprising carrier with T cell
PT epitope and many carbohydrate fragments with B cell epitopes,
PT particularly for vaccination against meningitis and diagnosis
XX
PS Example 6; Page 53; 83pp; English.
XX
CC The present invention describes a multivalent immunogenic molecule (I)
CC comprising: (i) carrier (Ia) having at least one functional T-cell
CC epitope; and (ii) many different carbohydrate fragments (Ib), all linked
CC to (Ia) and each having at least one functional B-cell epitope. (Ia)
CC increases the immunogenicity of (Ib). AAV29819 to AAV29826 represent
CC potential T-cell epitopes from meningococcal and pneumococcal proteins,
CC used in the exemplification of the present invention. (I) are used to
CC generate an immune response, specifically for protective vaccination
CC against meningitis (Streptococcus pneumoniae or Neisseria meningitidis),

CC but also against tumour-related antigens and antigens from other
CC bacteria, e.g. Escherichia coli, Salmonella typhi, Streptococcus mutans,
CC Cryptococcus neoformans, Klebsiella, Staphylococcus aureus or
CC Pseudomonas aeruginosa, to detect, by complex formation, (1)-reactive
CC antibodies and to raise (Ib)-specific antibodies, either for diagnostic
CC detection of the corresponding antigen in usual immunoassays or, if
CC directed against tumour antigens, for conjugation to anticancer agents.
CC The combination of T- and B-cell epitopes in a single vaccine results
CC in a strong and long-lasting humoral immunity.
XX
SQ Sequence 27 AA:
Query Match 5.2%; Score 15; DB 20; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 132 KNIAKQIAKDPKRN 146
DB 6 KNIAKQIAKDPKRN 20
|||||
RESULT 24
AAB07441
ID AAB07441 standard; Peptide; 13 AA.
XX
AC AAB07441;
XX
DT 20-OCT-2000 (first entry)
XX
DE Antigenic peptide from a lipoprotein of an ABC transporter protein.
XX
KW Lipoprotein; ATP-binding cassette transporter; MtsA; immune response;
KW vaccine.
XX
OS Streptococcus pyogenes.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /note- "this residue is not part of the native sequence"
FT
PN WO200040729-A1.
XX
PD 13-JUL-2000.
XX
PE 30-DEC-1999; 99WO-GB04445.
XX
PR 31-DEC-1998; 98GB-0028880.
XX
PA (ACTI-) ACTINOVA LTD.
XX
PI Bjoerck L, Janulczyk R;
XX
DR WPI; 2000-465989/40.
XX
PT New polypeptide comprising an ATP-binding cassette transporter of
PT Streptococcus pyogenes for use as a vaccine against Streptococcus and
PT for assays that detect immune reactivity to the polypeptides in animals
PT and humans -
XX
PS Example 7; Page 31; 55pp; English.
XX
CC AAB07437-41 represent antigenic peptides derived from a Streptococcus
CC pyogenes polypeptide which is lipoprotein of an ATP-binding cassette
CC transporter (MtsA). The polypeptide generates a protective immune
CC response to Streptococcus, preferably group A, such as S. pyogenes.
CC It is used to prepare a vaccine against Streptococcus. The new
CC polypeptides may be used in serological or cell mediated immune
CC assays for the detection of immune reactivity to the polypeptides in
CC animals and humans. Antibodies to MtsA protein can be detected using
CC an immunoassay with the polypeptides.
XX
SQ Sequence 13 AA:


Query Match 4.2%; Score 12; DB 21; Length 13;
 Best Local Similarity 100.0%; Pred. No. 8.7e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 44 DPHYEPLPED 55
 |||||||||
 DB 1 DPHYEPLPED 12

RESULT 25

AAV00023
 ID AAY00023 standard; Protein; 289 AA.

AC AAY00023;

DT 20-APR-1999 (first entry)

DE Enterococcus faecalis antigenic polypeptide fragment EF008.

KM Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
 detection; attenuation; antigenic.

OS Enterococcus faecalis.

PN WO9850554-A2.

PD 12-NOV-1998.

PF 04-MAY-1998; 98WO-US08959.

PR 14-NOV-1997; 97US-0066009.

PR 06-MAY-1997; 97US-0044031.

PR 16-MAY-1997; 97US-0046655.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Bailey C, Choi GH, Bromocky J A, Kunsch CA;

DR WPI; 1999-070095/06.

DR N-PSDB; AAX30013.

PT New isolated Enterococcus faecalis polynucleotides - used to develop

PT products for the detection of Enterococcus and for use in vaccines

PT for prevention or attenuation of Enterococcus infection

PS Claim 9; Page 86; 301pp; English.

XX The present sequence represents an antigenic polypeptide fragment

XX isolated from Enterococcus faecalis. The present invention describes

XX genes, proteins and antigenic polypeptides isolated from E. faecalis.

XX The proteins can be used in vaccines for preventing or attenuating an

XX infection caused by a member of the Enterococcus genus in an animal.

XX They can also be used for detecting Enterococcus antibodies in a sample.

XX The nucleotide sequences can be used for detecting Enterococcus nucleic

XX acids. Products from the present invention can also be used for

XX screening compounds to identify agonists and antagonists of E. faecalis

XX protein activity.

SO Sequence 289 AA;

Query Match 3.8%; Score 11; DB 20; Length 289;

Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 DPHYEPLPED 55
 |||||||||

DB 44 DPHYEPLPED 54

RESULT 26
 ABP43242
 ID ABP43242 standard; Protein; 289 AA.

AC ABP43242;

XX 05-AUG-2002 (first entry)

DE E faecalis EF008 antigenic fragment.

KM Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.

OS Enterococcus faecalis.

PN US2002045737-A1.

PD 18-APR-2002.

PF 04-MAY-1998; 98US-0071035.

PR 04-MAY-1998; 98US-0071035.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Choi GH, Bailey C, Bromocky J A, Kunsch CA;

DR WPI; 2002-425450/45.

DR N-PSDB; ABN97998.

PT New genes and polypeptides from Enterococcus faecalis, useful as

PT vaccines for preventing, treating or attenuating an infection caused by

PT a member of the Enterococcus genus in an animal, particularly E.

PT faecalis

PS Claim 9; Page 34-35; 255pp; English.

XX The present invention provides the protein and coding sequences of a

XX number of polypeptides from Enterococcus faecalis. The proteins can be

XX used as vaccines for preventing or attenuating an infection caused by a

XX member of the Enterococcus genus in an animal, particularly E. faecalis.

XX The polynucleotide is also useful for preventing or treating E. faecalis

XX infection. The present sequence is a protein of the invention.

SO Sequence 289 AA;

Query Match 3.8%; Score 11; DB 23; Length 289;

Best Local Similarity 100.0%; Pred. No. 0.015;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 DPHYEPLPED 55
 |||||||||

DB 44 DPHYEPLPED 54

RESULT 27

AAV00022

ID AAY00022 standard; Protein; 308 AA.

AC AAY00022;

DT 20-APR-1999 (first entry)

DE Enterococcus faecalis protein EF008.

KM Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
 detection; attenuation; antigenic.

OS Enterococcus faecalis.

PN WO9850554-A2.

PD 12-NOV-1998.

PF 04-MAY-1998; 98WO-US08959.

PR 14-NOV-1997; 97US-0066009.

PR 06-MAY-1997; 97US-0044031.

PR 16-MAY-1997; 97US-0046655.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
PI Bailey C, Choi GH, Hromocky J A, Kunsch CA;
XX WPI; 1999-070095/06.
DR N-PSDB; AAX20012.
XX
PT New isolated Enterococcus faecalis polynucleotides - used to develop
PT products for the detection of Enterococcus and for use in vaccines
PT for prevention or attenuation of Enterococcus infection
XX
PS Claim 9; Page 85; 301pp; English.
XX
CC The present sequence represents a protein isolated from
CC Enterococcus faecalis. The present invention describes genes, proteins
CC and antigenic polypeptides isolated from E. faecalis. The proteins can
CC be used in vaccines for preventing or attenuating an infection caused
CC by a member of the Enterococcus genus in an animal. They can also be
CC used for detecting Enterococcus antibodies in a sample. The nucleotide
CC sequences can be used for detecting Enterococcus nucleic acids.
CC Products from the present invention can also be used for screening
CC compounds to identify agonists and antagonists of E. faecalis protein
CC activity.
XX
SQ Sequence 308 AA;
XX
Query Match 3.8%; Score 11; DB 20; Length 308;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 45 DPHEYEPLPED 55
DB 63 DPHEYEPLPED 73
XX
RESULT 28
ABP43241
ID ABP43241 standard; Protein: 308 AA.
XX
AC ABP43241;
XX
DT 05-AUG-2002 (first entry)
XX
DE E faecalis EF008 protein.
XX
KW Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic.
XX
OS Enterococcus faecalis.
XX
PN US2002045737-A1.
XX
PD 18-APR-2002.
XX
PF 04-MAY-1998; 98US-0071035.
XX
PR 04-MAY-1998; 98US-0071035.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Choi GH, Bailey C, Hromocky J A, Kunsch CA;
XX WPI; 2002-425450/45.
DR N-PSDB; ABN97997.
XX
XX New genes and polypeptides from Enterococcus faecalis, useful as
PT vaccines for preventing, treating or attenuating an infection caused by
PT a member of the Enterococcus genus in an animal, particularly E.
PT faecalis
XX
PS Claim 9; Page 34; 255pp; English.
XX

CC The present invention provides the protein and coding sequences of a
CC number of polypeptides from Enterococcus faecalis. The proteins can be
CC used as vaccines for preventing or attenuating an infection caused by a
CC member of the Enterococcus genus in an animal, particularly E. faecalis.
CC The polynucleotide is also useful for preventing or treating E. faecalis
CC infection. The present sequence is a protein of the invention.
XX
SQ Sequence 308 AA;
XX
Query Match 3.8%; Score 11; DB 23; Length 308;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 45 DPHEYEPLPED 55
DB 63 DPHEYEPLPED 73
XX
RESULT 29
ABBA9144
ID ABBA9144 standard; Protein: 310 AA.
XX
AC ABBA9144;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #1848.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussauget O, Chetoui F, Medjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Madueno E, De Pablos B, Wehlund J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX
DR WPI; 2002-010914/01.
XX
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides
XX
PS Claim 6; SEQ ID No 1849; 192pp; French.
XX
XX The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and

CC vaccines compositions for the treatment or prevention of infections by L.
 CC monocytes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.

CC Sequence 310 AA;

Query Match 3.1%; Score 9; DB 23; Length 310;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 200 AYIMEINTE 208
 |||||
 Db 221 AYIMEINTE 229

RESULT 30
 AAU46618
 ID AAU46618 standard; Protein; 54 AA.

XX AAU46618:

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #7514.

XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX MO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59534.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -

XX Example 1; SEQ ID NO 7813; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as

CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.

CC Sequence 54 AA;

Query Match 2.8%; Score 8; DB 22; Length 54;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 LIYVSECC 187
 |||||
 Db 23 LIYVSECC 30

RESULT 31
 ABP04023
 ID ABP04023 standard; Protein; 196 AA.

XX ABP04023:

XX 24-JUN-2002 (first entry)

XX Human ORFX protein sequence SEQ ID NO:8028.

XX Human: open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.

XX Homo sapiens.

XX WO200192523-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US10836.

XX 30-MAY-2000; 2000US-206132P.

XX 29-AUG-2000; 2000US-228716P.

XX (CURA-) CURAGEN CORP.

XX Shinketsu RA, Leach MD;

XX WPI; 2002-106308/14.

XX N-PSDB; ABN19775.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders -

XX Disclosure; SEQ ID 8028; 1037pp; English.

XX The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification)). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester

storage disease, various immune deficiencies and disorders, infectious diseases, autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. ORF proteins are also useful for treating burns, incisions, ulcers, for treating osteoporosis, bone degenerative disorders, or periodontal disease, and for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues and conditions resulting from systemic cytokine damage.

N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 196 AA;

Query Match 2.8%; Score 8; DB 23; Length 196;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 NLENGIIV 130
| | | | |
Db 15 NLENGIIV 22

RESULT 32

AAU36295
ID AAU36295 standard; Protein: 255 AA.

AC AAU36295;

DT 14-FEB-2002 (first entry)

DE Pseudomonas aeruginosa cellular proliferation protein #285.

KW Antisense: prokaryotic cellular proliferation protein;

KM antibiotic; antibacterial; drug design.

OS Pseudomonas aeruginosa.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

DR WPI: 2001-611495/70.

DR N-PSDB; AAS54154.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

PS Example 3; Seq ID No 11888; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

Sequence 255 AA;

Query Match 2.8%; Score 8; DB 22; Length 255;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 208 EEEGPPDQ 215
| | | | |
Db 229 EEEGPPDQ 236

RESULT 33

ABB70862
ID ABB70862 standard; Protein: 269 AA.

AC ABB70862;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 39378.

KW Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EW;

DR WPI: 2001-656860/75.

DR N-PSDB; ABL14965.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PS interactions -

XX Disclosure; SEQ ID NO 39378; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (AB57737-AB572072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 269 AA;

2.8%; Score 8; DB 22; Length 269;

Query Match
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 EPLPEDAE 57
 |||||
 Db 58 EPLPEDAE 65

RESULT 34

AAW26367

ID AAW26367 standard; Protein: 309 AA.

AAW26367;

03-DEC-1997 (first entry)

Staphylococcus aureus saliva binding protein.

Saliva binding protein; vaccine; diagnosis; gene therapy; antibacterial; genetic immunisation; antibody; antagonist.

Staphylococcus aureus strain WCUH 29 (NCIMB 40771).

W09714800-A1.

24-APR-1997.

15-OCT-1996; 96WO-GB02530.

01-AUG-1996; 96GB-0016136.

16-OCT-1995; 95GB-0021147.

04-MAR-1996; 96GB-0004599.

(SMK) SMITHKLINE BEECHAM PLC.

Burnham MKR, Hodgson JE;

WPI: 1997-245116/22.

N-PSDB; AAT84529.

New nucleic acid encoding new saliva binding protein of

Staphylococcus aureus - useful for vaccination against and treatment

of bacterial infections, optionally by gene therapy

Claim 1; Page 34-35; 48pp; English.

This protein sequence comprises a novel saliva binding protein of Staphylococcus aureus WCUH 29 that is useful as an antibacterial agent or vaccine. Its sequence was deduced from a polynucleotide (AAT84529) isolated from a WCUH 29 chromosomal library. Also claimed are: a vector containing the DNA; host cells containing the vector; an antibody (Ab) against saliva binding protein; antagonists that inhibit the activity of saliva binding protein; a method for identifying compounds that bind to, and inhibit activity of, the saliva binding protein; and a method for expressing saliva binding protein in vivo, through gene therapy, as a means of inducing an immunological response in a mammal. The polypeptide and its inhibitors can be used to prevent adhesion of Gram-positive bacteria to extracellular matrix proteins on in-dwelling devices, in wounds etc., and to block cell invasion mediated by cell surface proteins. They can also be used for pre-operative protection, before dental treatment or generally as wound treatment.

SQ Sequence 309 AA;

2.8%; Score 8; DB 18; Length 309;

Query Match
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GODPHEYE 50
 |||||
 Db 63 GODPHEYE 70

RESULT 35

ABP40702

ID ABP40702 standard; Protein: 316 AA.

ABP40702;

24-JUL-2002 (first entry)

Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5547.

Staphylococcus epidermidis; open reading frame; ORF; bacterial infection; antibacterial; gene therapy.

Staphylococcus epidermidis.

US6380370-B1.

30-APR-2002.

13-AUG-1998; 98US-0134001.

14-AUG-1997; 97US-055779P.

08-NOV-1997; 97US-064964P.

(GENO-) GENOME THERAPEUTICS CORP.

Doucette-Stamm LA, Bush D;

WPI: 2002-381255/41.

N-PSDB; ABN93247.

Novel isolated nucleic acid encoding a Staphylococcus epidermidis

polypeptide, useful for diagnosing and treating bacterial infections -

Disclosure; SEQ ID 5547; 267pp; English.

ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading

frame (ORF) nucleic acid sequences which encode the amino acid sequences

given in ABP35124 to ABP37960. The S. epidermidis sequences have

antibacterial activity and can be used in gene therapy. The sequences

can also be used in the diagnosis and treatment of bacterial infections,

particularly S. epidermidis infections. The sequences can be used to

screen for compounds able to interfere with the S. epidermidis life

cycle or inhibit S. epidermidis infection.

N.B. The sequence data for this patent did not form part of the printed

specification, but was obtained in electronic format directly from the

USPTO web site.

SQ Sequence 316 AA;

2.8%; Score 8; DB 23; Length 316;

Query Match
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GODPHEYE 50
 |||||
 Db 70 GODPHEYE 77

RESULT 36

AAP71178

ID AAP71178 standard; Protein: 60 AA.

AAP71178;

15-APR-1991 (first entry)

Bacillus exoprotease secretion promoting protein.

XX XX Exoprotease.
KW JF61282400-A.
XX 12-DEC-1986.
XX 07-JUN-1985; 85JP-0123736.
XX 07-JUN-1985; 85JP-0123736.
XX (MITU) MITSUBISHI CHEM IND KR.
XX WPI: 1987-025820/04.
XX N-PSDB: AAN71310.
XX Protein promoting extra:cellular prodn. of exo:protease - by use
PT of recombinant DNA technique
XX Claim 3; Page 1060; 8pp; Japanese.
XX Protein may be produced from a bacterial expression vector, and has
CC extracellular exoprotease production promoting activity.
XX Sequence 60 AA;
SQ

Query Match 2.4%; Score 7; DB 8; Length 60;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 IAKKGKP 268
Db 27 IAKKGKP 33
|||||

RESULT 37
AAG20210
ID AAG20210 standard; Protein; 98 AA.
XX
XX AAG20210;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 22310.
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132207.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147436.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147933.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151338.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 15-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159329.
 PR 13-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.

PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 22-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161982.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 2.4%; Score 7; DB 21; Length 98;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSTGAKT 7
 Db 6 SSTGAKT 12

RESULT 38
 ABP34948 standard; Protein; 98 AA.
 ID ABP34948;
 AC ABP34948;
 XX 08-JUL-2002 (First entry)

DE Human ORF3921 protein, SEQ ID NO:7842.
 XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; actinin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infectious; neoplastic;
 KW vasotrophic; antipsoriatic; antidiabetic; cytostatic; thrombolytic;
 KW neuroprotective; antithrombotic; antihypertensive; thrombolytic;
 KW cardiact; hypotensive; antihypertensive; antihypertensive; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.

OS Homo sapiens.
 PN WO200190366-A2.
 XX 29-NOV-2001.
 PD 24-MAY-2001; 2001WO-US17076.
 XX 24-MAY-2000; 2000US-206690P.
 PF (CURA-) CURAGEN CORP.
 XX Leach MD, Shinkets RA;
 PI WPI. 2002-106200/14.
 DR N-PSDB; ABR78574.
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation

XX Claim 10; Page 2206; 2508pp; English.

CC Sequences ABP1028-ABP3561 represent 4534 novel human proteins
CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
CC ABN79587 represent cDNAs encoding them. The invention also encompasses
CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
CC polynucleotides, the recombinant production of ORFX proteins, antibodies
CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
CC polypeptides, methods of screening for modulators of ORFX expression or
CC activity, and methods of screening individuals for a predisposition to an
CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antiinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases.

CC SQ Sequence 98 AA;

Query Match 2.4%; Score 7; DR 23; Length 98;

Best Local Similarity 100.0%; Pred. No. 61;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 263 AKKGKPG 269

DB 18 AKKGKPG 24

RESULT 39

AAG20209

ID AAG20209 standard; Protein; 109 AA.

XX AAG20209;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 22309.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX Arabidopsis thaliana.

PN Arabidopsis thaliana.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

```

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148368.
PR 17-AUG-1999; 99US-0148175.
PR 18-AUG-1999; 99US-0148426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151938.
PR 01-SEP-1999; 99US-0152263.
PR 07-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

```

```

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 2.4%; Score 7; DB 21; Length 109;
Best local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SGTAKT 7
DB 17 SGTAKT 23

```

```

RESULT 40
ABB52551
ID ABB52551 standard; Protein; 135 AA.
AC ABB52551;
XX

```

```

DT 11-FEB-2002 (first entry)
XX

```

```

DE Escherichia coli polypeptide SEQ ID NO 471.
XX

```

```

XX Escherichia coli; B2/DNA-; antiinflammatory; antibacterial;
XX immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
XX systemic infection; non-diarrhoeal infection; septicaemia;
XX pyelonephritis; antibiotic resistance.
XX

```

```

OS Escherichia coli.
XX

```

```

XX WO20016572-A2.
XX

```

```

XX 13-SEP-2001.
XX

```

```

XX 12-MAR-2001; 2001WO-EP03445.
XX

```

```

XX 10-MAR-2000; 2000EP-0003145.
XX

```

```

XX 02-FEB-2001; 2001FR-0001449.
XX

```

```

XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX

```

```

XX Bingen E, Bonaccorsl S, Clermont O, Nassif X, Tinsley C;
XX WPI; 2001-550253/61.
DR

```

XX A library of DNA fragments of *Escherichia coli* strains for the
 PT phylogenetic determination of a given strain comprises polynucleotides of
 PT nature B2/D+ A- -
 PS Example 6; Fig 6; 646bp; English.
 XX The invention relates to a library of DNA fragments of *Escherichia coli*
 CC strains comprising polynucleotides (ABA8577-ABA88729 and ABA89533)
 CC and encoded proteins (ABBS2459-ABBS2919 and ABBS2954-ABBS3094) of nature
 CC B2/D+A-. The polynucleotides have potential anti-inflammatory,
 CC antibacterial and immunosuppressive activity as part of pharmaceutical
 CC compositions used to treat, palliate or prevent extra-intestinal *E. coli*
 CC infections. The polypeptides are useful for determining the phylogenetic
 CC group of a given *E. coli* strain. These polypeptides can detect and treat
 CC an undesired development of *E. coli*, particularly an extra-intestinal
 CC infection that include systemic and non-diarrhoeal infections such as
 CC septicaemia, pyelonephritis and meningitis this is particularly
 CC advantageous as bacterial resistance is increasing with the more
 CC frequent use of broad spectrum antibiotics.
 XX Sequence 135 AA;
 SQ
 Query Match 2.4%; Score 7; DB 22; Length 135;
 Best Local Similarity 100.0%; Pred. No. 81;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 157 VAKLEKL 163
 Db 44 VAKLEKL 50
 RESULT 41
 AAW14554
 ID AAW14554 standard; Protein; 142 AA.
 XX
 AC AAW14554;
 XX
 DT 28-OCT-1997 (first entry)
 XX
 DE Streptococcus pneumoniae PspA N-terminal end.
 XX
 KW PspA; pneumococcal surface protein; vaccine; otitis media;
 KW meningitis; bacteraemia; pneumonia.
 XX
 OS Streptococcus pneumoniae strain Bg6692pro.
 XX
 PN WO9709994-A1.
 PD 20-MAR-1997.
 XX
 PF 16-SEP-1996; 96WO-US14819.
 XX
 PR 15-SEP-1995; 95US-0529055.
 XX
 PA (UABR-) UAB RES FOUND.
 XX
 PI Briles DE, Brooks-Walter A, Crain MJ, Hollingshead S;
 PI McDaniel US, Swiatlo E, Tart R, Yother J;
 DR WPI: 1997-202002/18.
 XX
 PT Streptococcus pneumoniae surface protein PspC and truncated PspA -
 PT used in vaccines for protecting animals against *S. pneumoniae*
 PT infection
 XX
 PS Example 6; Fig 13; 296bp; English.
 XX
 CC This sequence shows the N-terminal end of pneumococcal surface
 CC protein A (PspA) of Streptococcus pneumoniae strain Bg6692pro.
 CC Comparison of the N-terminal and central regions (AAW14533-57 and
 CC AAW14562-91) of PspA polypeptides from different pneumococcal strains
 CC can be used to divide the strains into several families based on

CC sequence homologues. PspA polypeptides, or fragments of them, can
 CC be used in vaccines to protect animals against *S. pneumoniae*.
 CC infection and hence for the prevention of diseases such as otitis
 CC media, meningitis, bacteraemia and pneumonia. The sequence of the
 CC 3' half of the PspA alpha-helical region and the immediate 5' tip
 CC of the coding sequence are likely to be the critical sequences for
 CC predicting PspA cross-reactions and vaccine composition.
 XX
 SQ Sequence 142 AA;
 Query Match 2.4%; Score 7; DB 18; Length 142;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 178 KRLIYVS 184
 Db 4 KRLIYVS 10
 RESULT 42
 AAG81390
 ID AAG81390 standard; Protein; 166 AA.
 XX
 AC AAG81390;
 XX
 DT 10-SEP-2001 (first entry)
 XX
 DE Human AFP protein sequence SEQ ID NO:298.
 XX
 KW Human; secreted protein; secretion; bacterial cell; fungal cell;
 KW eukaryotic cell; fusion protein; maltose binding protein;
 KW immunoglobulin constant region; polystyridine tag.
 XX
 OS Homo sapiens.
 XX
 PN WO200129221-A2.
 XX
 PD 26-APR-2001.
 XX
 PF 20-OCT-2000; 2000WO-US29052.
 XX
 PR 20-OCT-1999; 99US-0160712.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Conklin DC, Yee DP;
 PI
 DR WPI: 2001-300340/31.
 DR N-PSDB; AAH52241.
 XX
 PT Isolated polypeptide for directing secretion of proteins of interest
 PT from a host cell including, e.g. bacteria, includes contiguous amino
 PT acid residues of polypeptide with specified amino acids
 XX
 PS Claim 1; Page 509; 617pp; English.
 XX
 CC AAH52093 to AAH52303 encode the human secreted proteins given in AAG81242
 CC to AAG81453. The secreted proteins can be used for directing the
 CC secretion of proteins of interest from a host cell including bacteria,
 CC fungal cells, and cultured higher eukaryotic cells. The present invention
 CC also describes fusion proteins, where a secreted protein of the invention
 CC is operably linked via a peptide bond or peptide linker to a second
 CC protein selected from the group consisting of maltose binding protein,
 CC an immunoglobulin constant region, a polystyridine tag and a peptide
 CC given in AAG81453.
 XX
 SQ Sequence 166 AA;
 Query Match 2.4%; Score 7; DB 22; Length 166;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 158 AKLEKLD 164

Db 49 AKLEKID 55
|||||
RESULT 43
ABBB69922 standard; Protein; 184 AA.
ID ABB69922;
AC ABB69922;
XX 26-MAR-2002 (first entry)
DT 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 36558.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX W0200171042-A2.
FN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001MO-US09231.
PE
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR N-PSDB; ABL14025.
DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -
XX
XX Disclosure; SEQ ID NO 36558; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 184 AA;
SQ
Query Match 2.4%; Score 7; DB 22; Length 184;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 143 PKNKETY 149
| | | | |
DB 50 PKNKETY 56
RESULT 44
AAG11073
ID AAG11073 standard; Protein; 188 AA.
AC AAG11073;
XX
XX 17-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SEQ ID NO: 9653.
DE

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
FN
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
PE
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.4%: Score 7; DB 21; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 KIDKEAK 168
|||||
Db 119 KIDKEAK 125

RESULT 45
AAG50188
ID AAG50188 standard; Protein: 188 AA.
XX
AC AAG50188;
XX
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63572.
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX

PN EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121875.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0143777.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148368.
PR 17-AUG-1999; 99US-0148175.
PR 18-AUG-1999; 99US-0148426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159285.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.4%; Score 7; DB 21; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 162 KLDREAR 168
|||||||
Db 119 KLDREAR 125

Search completed: May 21, 2003, 11:11:55
Job time : 39 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 11:12:01 ; Search time 24 Seconds

(Without alignments)
1185.940 Million cell updates/sec

Title: US-09-869-677A-2

Perfect score: 287
Sequence: 1 SSGGAKTAKSDKLVATNS.....PGDSYYAMKMKNDKISEGL 287

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 375593 seqs, 99172665 residues

Word size : 0

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	9.4	289	US-09-765-272-20	Sequence 156, Appl
2	27	9.4	309	US-09-769-787-156	Sequence 20, Appl
3	15	5.2	27	US-09-027-956-7	Sequence 78, Appl
4	8	2.8	255	US-09-815-242-11888	Sequence 11888, A
5	7	2.4	194	US-09-834-765-768	Sequence 768, App
6	7	2.4	207	US-09-975-719-211	Sequence 211, App
7	7	2.4	239	US-10-280-794-7	Sequence 7, Appl
8	7	2.4	355	US-09-815-242-10739	Sequence 10739, A
9	7	2.4	577	US-09-815-242-11347	Sequence 11347, A
10	7	2.4	627	US-09-925-300-1516	Sequence 1516, A
11	7	2.4	724	US-10-211-962-74	Sequence 78, Appl
12	7	2.4	802	US-10-245-103-78	Sequence 78, Appl
13	7	2.4	802	US-10-245-107-78	Sequence 78, Appl
14	7	2.4	802	US-10-245-143-78	Sequence 78, Appl
15	7	2.4	802	US-10-245-771-78	Sequence 78, Appl
16	7	2.4	802	US-10-245-851-78	Sequence 78, Appl
17	7	2.4	802	US-10-245-883-78	Sequence 78, Appl
18	7	2.4	802	US-10-237-535-78	Sequence 78, Appl
19	7	2.4	802	US-10-238-183-78	Sequence 78, Appl

20	7	2.4	802	US-10-238-283-78	Sequence 78, Appl
21	7	2.4	802	US-10-238-370-78	Sequence 78, Appl
22	7	2.4	802	US-10-248-055-78	Sequence 78, Appl
23	7	2.4	802	US-10-245-147-78	Sequence 78, Appl
24	7	2.4	802	US-10-245-730-78	Sequence 78, Appl
25	7	2.4	802	US-10-245-739-78	Sequence 78, Appl
26	7	2.4	802	US-10-246-210-78	Sequence 78, Appl
27	7	2.4	802	US-10-239-196-78	Sequence 78, Appl
28	7	2.4	802	US-10-243-024-78	Sequence 78, Appl
29	7	2.4	802	US-10-243-409-78	Sequence 78, Appl
30	7	2.4	802	US-10-245-033-78	Sequence 78, Appl
31	7	2.4	802	US-10-245-621-78	Sequence 78, Appl
32	7	2.4	802	US-10-245-880-78	Sequence 78, Appl
33	7	2.4	802	US-10-243-095-78	Sequence 78, Appl
34	7	2.4	802	US-10-243-185-78	Sequence 78, Appl
35	7	2.4	802	US-10-245-427-78	Sequence 78, Appl
36	7	2.4	802	US-10-245-473-78	Sequence 78, Appl
37	7	2.4	802	US-10-245-770-78	Sequence 78, Appl
38	7	2.4	802	US-10-245-877-78	Sequence 78, Appl
39	7	2.4	802	US-10-246-976-78	Sequence 78, Appl
40	7	2.4	802	US-10-243-320-78	Sequence 78, Appl
41	6	2.1	23	US-09-853-253-5	Sequence 3, Appl
42	6	2.1	23	US-09-853-253-6	Sequence 6, Appl
43	6	2.1	24	US-09-853-253-4	Sequence 4, Appl
44	6	2.1	35	US-10-140-164-34	Sequence 34, Appl
45	6	2.1	35	US-10-140-164-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-09-765-272-20
Sequence 156, Appl
Patent No. US20020061545A1
GENERAL INFORMATION:
APPLICANT: Chol et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB34002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-09-765-272-20

Query Match 9.4%; Score 27; DB 10; Length 289;
Best Local Similarity 100.0%; Pred. No. 5.1e-18;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 IAGDKIDLSHIVPIGDPHEXEPLEPED 55
DB 29 IAGDKIDLSHIVPIGDPHEXEPLEPED 55

RESULT 2

US-09-769-787-156
; Sequence 156, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: FWC/P21129W0
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 156
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-156

Query Match 9.4%; Score 27; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 5.4e-18;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 IAGDKIDLSHIVPIGDPHEXEPLEPED 55
DB 49 IAGDKIDLSHIVPIGDPHEXEPLEPED 75

RESULT 3

US-09-027-956-7
; Sequence 7, Application US/09027956A
; Patent No. US20010048929A1
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: LINDBERG, Alf
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: NOVEL MULTI-OLIGOSACCHARIDE GLYCOCOMUNUGATE BACTERIAL
; FILE REFERENCE: 1038-791 MTS:1b
; CURRENT APPLICATION NUMBER: US/09/027,956A
; CURRENT FILING DATE: 1998-02-23
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-027-956-7

Query Match 5.2%; Score 15; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 KNIKOLIAADPKNK 146
DB 6 KNIKOLIAADPKNK 20

RESULT 4

US-09-815-242-11888
; Sequence 11888, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11888
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11888

Query Match 2.8%; Score 8; DB 10; Length 255;
Best Local Similarity 100.0%; Pred. No. 9.6;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 EEEGTPDQ 215
DB 229 EEEGTPDQ 236

RESULT 5

US-09-834-765-768
; Sequence 768, Application US/09834765
; Patent No. US20020055478A1
; GENERAL INFORMATION:
; APPLICANT: Mary Faris
; APPLICANT: Pia M. Chalilata-Eld
; APPLICANT: Arthur B. Raitano
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Daniel E. H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
; FILE REFERENCE: 129.60S01
; CURRENT APPLICATION NUMBER: US/09/834,765
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/197,647
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 770
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 768
; LENGTH: 194
; TYPE: PRT

ORGANISM: C.elegans
US-09-834-765-768

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 194;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SDKLKV 16
Db 74 SDKLKV 80

RESULT 6

US-09-975-719-211
Sequence 211, Application US/09975719
Publication No. US20030022349A1
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick M.
APPLICANT: Rahme, Laurence G.
TITLE OF INVENTION: VIRUS-ASSOCIATED NUCLEIC ACID
FILE REFERENCE: 00786/361003
CURRENT APPLICATION NUMBER: US/09/975,719
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 09/199,637
PRIOR FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: US 60/066,517
PRIOR FILING DATE: 1997-11-25
NUMBER OF SEQ ID NOS: 437
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 211
LENGTH: 207
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-975-719-211

Query Match
Best Local Similarity 100.0%; Score 7; DB 9; Length 207;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ATNSIIA 23
Db 149 ATNSIIA 155

RESULT 7

US-10-280-794-7
Sequence 7, Application US/10280794
Publication No. US20030087391A1
GENERAL INFORMATION:
APPLICANT: Bandman, Olga; Hillman, Jennifer L.
APPLICANT: Goll, Surya K.; Capitant, Sherry A.
TITLE OF INVENTION: B CELL RECEPTOR ASSOCIATED PROTEINS
FILE REFERENCE: PE-0188-1 DIV
CURRENT APPLICATION NUMBER: US/10/280,794
CURRENT FILING DATE: 1997-05-07
PRIOR APPLICATION NUMBER: 08/851,971
PRIOR FILING DATE: 1997-05-07
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 239
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: GenBank ID No. US20030087391A1 g541730
US-10-280-794-7

Query Match
Best Local Similarity 100.0%; Score 7; DB 9; Length 239;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 AENKKLI 181
Db 173 AENKKLI 179

RESULT 8

US-09-815-242-10739
Sequence 10739, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10739
LENGTH: 355
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10739

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 355;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 KIKVVAT 18
Db 237 KIKVVAT 243

RESULT 9

US-09-815-242-11347
Sequence 11347, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078

;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 11347
;; LENGTH: 577
;; TYPE: PRT
;; ORGANISM: Helicobacter pylori
US-09-815-242-11347

Query Match 2.4%; Score 7; DB 10; Length 577;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 AKDPKPK 146
Db 311 AKDPKPK 317

RESULT 10
;; Sequence 1516, Application US/09925300
;; Patent No. US20020151681A1
;; GENERAL INFORMATION:
;; APPLICANT: Craig Rosen,
;; APPLICANT: Steve Ruben,
;; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;; FILE REFERENCE: PA101
;; CURRENT APPLICATION NUMBER: US/09/925,300
;; CURRENT FILING DATE: 2001-08-10
;; PRIOR APPLICATION NUMBER: PCT/US00/05988
;; PRIOR FILING DATE: 2000-03-08
;; PRIOR APPLICATION NUMBER: 60/124,270
;; PRIOR FILING DATE: 1999-03-12
;; NUMBER OF SEQ ID NOS: 1890
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 1516
;; LENGTH: 627
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-925-300-1516

Query Match 2.4%; Score 7; DB 10; Length 627;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 YGVPSAY 201
Db 318 YGVPSAY 324

RESULT 11
;; Sequence 24, Application US/10211962
;; Publication No. US20030082640A1
;; GENERAL INFORMATION:
;; APPLICANT: Herz, Joachim
;; APPLICANT: Gotthardt, Michael
;; TITLE OF INVENTION: LDL Receptor Signaling Pathways
;; FILE REFERENCE: UTSW0708
;; CURRENT APPLICATION NUMBER: US/10/211,962
;; CURRENT FILING DATE: 2002-08-01

;; PRIOR APPLICATION NUMBER: US/09/562,737
;; PRIOR FILING DATE: 2000-05-01
;; NUMBER OF SEQ ID NOS: 132
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 24
;; LENGTH: 724
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-211-962-24

Query Match 2.4%; Score 7; DB 9; Length 724;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 LKVIKPS 229
Db 241 LKVIKPS 247

RESULT 12
;; Sequence 78, Application US/10245103
;; Publication No. US20030068778A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin
;; APPLICANT: Bacon, Dan
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Philippe
;; APPLICANT: Watande, Colin
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; APPLICANT: Fong, Sherman
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P9630R1C112
;; CURRENT APPLICATION NUMBER: US/10/245,103
;; CURRENT FILING DATE: 2002-09-17
;; PRIOR APPLICATION NUMBER: 10/197942
;; PRIOR FILING DATE: 2002-07-18
;; PRIOR APPLICATION NUMBER: 60/059114
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/063046
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/065027
;; PRIOR FILING DATE: 1997-11-10
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/086478
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090689
;; PRIOR FILING DATE: 1998-06-25
;; Remaining Prior Application data removed - See File Wrapper or PAM.
;; NUMBER OF SEQ ID NOS: 116
;; SEQ ID NO 78
;; LENGTH: 802
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-245-103-78

Query Match 2.4%; Score 7; DB 9; Length 802;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 82 KLVKNAQ 88
Db 413 KLVKNAQ 419

RESULT 13

US-10-245-107-78
; Sequence 78, Application US/10245107
; Publication No. US2003006879A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C71
; CURRENT APPLICATION NUMBER: US/10/245,107
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-107-78

Query Match 2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 82 KLVKNAQ 88
Db 413 KLVKNAQ 419

RESULT 14
US-10-245-143-78
; Sequence 78, Application US/10245143
; Publication No. US20030068780A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin

; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C90
; CURRENT APPLICATION NUMBER: US/10/245,143
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-143-78

Query Match 2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 82 KLVKNAQ 88
Db 413 KLVKNAQ 419

RESULT 15
US-10-245-771-78
; Sequence 78, Application US/10245771
; Publication No. US20030068781A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C98

```

; CURRENT APPLICATION NUMBER: US/10/245,771
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See file wrapper or PAM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-771-78
```

```

Query Match          2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      82 KLVKNAQ 88
        |||||
Db       413 KLVKNAQ 419
```

```

RESULT 16
; Sequence 78, Application US/10245851
; Publication No. US20030068782A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Collin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C93
; CURRENT APPLICATION NUMBER: US/10/245,851
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
```

```

; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See file wrapper or PAM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-851-78
```

```

Query Match          2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      82 KLVKNAQ 88
        |||||
Db       413 KLVKNAQ 419
```

```

RESULT 17
; Sequence 78, Application US/10245883
; Publication No. US20030068783A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Collin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C70
; CURRENT APPLICATION NUMBER: US/10/245,883
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See file wrapper or PAM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
```

US-10-245-883-78

Query Match 2.4% Score 7: DB 9: Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 KLVKNAQ 88
|||||
DB 413 KLVKNAQ 419

RESULT 18
US-10-237-535-78

Sequence 78, Application US/10237535
Publication No. US20030073188A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Phillippe
APPLICANT: Matande, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3630R1C3
CURRENT FILING DATE: 2002-09-06
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091358
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/106932
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/115554
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119342
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/123957
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/127372
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/131271
PRIOR FILING DATE: 1999-04-27

PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/135725
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135729
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/136385
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144732
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144790
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145228
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146843
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/148188
PRIOR FILING DATE: 1999-08-10
PRIOR APPLICATION NUMBER: 60/148513
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/149327
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149395
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/150114
PRIOR FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: 60/151700
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/151734
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/177118
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: 60/179851
PRIOR FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 60/180921
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/198587
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 60/199614
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 60/206330
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206368
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/209932
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/218371
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: 60/222695
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/229896
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/230621
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/232887

PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/235147
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/261878
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/261910
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/261939
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/262150
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/264395
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/266421
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/267623
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/274399
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/280982
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/282129
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/282129
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/290589
PRIOR FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 09/267213
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/924419
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/927796
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/929404
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 09/931836
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 09/941992
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 09/946374
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 10/001054
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 10/05586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 10/081056
PRIOR FILING DATE: 2002-02-20
PRIOR APPLICATION NUMBER: 10/119480
PRIOR FILING DATE: 2002-04-09

Query Match 2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

OY 82 KLVKNAQ 88

Db 413 KLVKNAQ 419
|||||||
RESULT 19
US-10-238-183-78
Sequence 78, Application US/10238183
Publication No. US20030073189A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Watanabe, Collin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3630R1C11
CURRENT APPLICATION NUMBER: US/10/238,183
PRIOR FILING DATE: 2002-09-09
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091358
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/106932
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/115554
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119342
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/123957
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/127372
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/131271
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/135725
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135729
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135750

PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138385
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144732
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144790
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145228
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146843
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/148188
PRIOR FILING DATE: 1999-08-10
PRIOR APPLICATION NUMBER: 60/148513
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/149327
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149395
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/150114
PRIOR FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: 60/151700
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/151734
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/177118
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: 60/179851
PRIOR FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 60/180921
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/198587
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 60/199614
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 60/206330
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206368
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/209832
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/218371
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: 60/222695
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/222896
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/230621
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/235147
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/261878
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/261910
PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: 60/261939
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/262150
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/264395
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/266421
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/267623
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/274399
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/280982
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/282129
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/282199
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/290589
PRIOR FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: 60/290997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 60/267213
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 60/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 60/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 60/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 60/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/924419
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 60/927796
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/929404
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 60/931836
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 60/941992
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 60/946374
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 60/001054
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/081056
PRIOR FILING DATE: 2002-02-20
PRIOR APPLICATION NUMBER: 60/119480
PRIOR FILING DATE: 2002-04-09

Query Match

2.4% Score 7; DB 9; Length 802;

Best Local Similarity 100.0%; Pred. No. 2.5e+02; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

82 KLVKNAO 88
|||||||

DB

413 KLVKNAQ 419

RESULT 20

US-10-238-283-78

; Sequence 78, Application US/10238283

```
Publication NO. US20030073190A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Watanabe, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P3630R1C15
CURRENT FILING DATE: 2002-09-09
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 78
LENGTH: 802
TYPE: PRT
ORGANISM: Homo Sapien
US-10-238-283-78

Query Match          2.4%; Score 7; DB 9; Length 802;
Best local similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      82 KLVKNAQ 88
DB      413 KLVKNAQ 419
```

```
RESULT 21
US-10-238-370-78
Sequence 78, Application US/10238370
Publication NO. US20030073191A1
```

```
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Watanabe, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
```

```
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P3630R1C10
CURRENT FILING DATE: 2002-09-09
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 78
LENGTH: 802
TYPE: PRT
ORGANISM: Homo Sapien
US-10-238-370-78

Query Match          2.4%; Score 7; DB 9; Length 802;
Best local similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      82 KLVKNAQ 88
DB      413 KLVKNAQ 419
```

```
RESULT 22
US-10-245-055-78
Sequence 78, Application US/10245055
Publication NO. US20030073192A1
```

```
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Watanabe, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P3630R1C88
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
```

;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/086478
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090689
;; PRIOR FILING DATE: 1998-06-25
;; Remaining Prior Application data removed - See file Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 116
;; LENGTH: 802
;; SEQ ID NO 78
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-245-055-78

Query Match 2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 82 KLVKNAQ 88
|||||
DB 413 KLVKNAQ 419

RESULT 23
US-10-245-147-78

;; Sequence 78, Application US/10245147
;; Publication No. US20030073193A1
;; GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin
;; APPLICANT: Baton, Dan
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Phillippe
;; APPLICANT: Watanabe, Colin
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; APPLICANT: Fong, Sherman
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; TITLE OF INVENTION: ACIDS ENCODING THE SAME
;; FILE REFERENCE: P3630R1C72
;; CURRENT APPLICATION NUMBER: US/10/245,147
;; CURRENT FILING DATE: 2002-09-16
;; PRIOR APPLICATION NUMBER: 10/197942
;; PRIOR FILING DATE: 2002-07-18
;; PRIOR APPLICATION NUMBER: 60/059114
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/063046
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/065027
;; PRIOR FILING DATE: 1997-11-10
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/086478
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090689
;; PRIOR FILING DATE: 1998-06-25
;; Remaining Prior Application data removed - See file Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 116
;; SEQ ID NO 78

;; LENGTH: 802
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-245-147-78

Query Match 2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 82 KLVKNAQ 88
|||||
DB 413 KLVKNAQ 419

RESULT 24
US-10-245-730-78

;; Sequence 78, Application US/10245730
;; Publication No. US20030073194A1
;; GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin
;; APPLICANT: Baton, Dan
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Phillippe
;; APPLICANT: Watanabe, Colin
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; APPLICANT: Fong, Sherman
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; TITLE OF INVENTION: ACIDS ENCODING THE SAME
;; FILE REFERENCE: P3630R1C85
;; CURRENT APPLICATION NUMBER: US/10/245,730
;; CURRENT FILING DATE: 2002-09-16
;; PRIOR APPLICATION NUMBER: 10/197942
;; PRIOR FILING DATE: 2002-07-18
;; PRIOR APPLICATION NUMBER: 60/059114
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/063046
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/065027
;; PRIOR FILING DATE: 1997-11-10
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/086478
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090689
;; PRIOR FILING DATE: 1998-06-25
;; Remaining Prior Application data removed - See file Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 116
;; SEQ ID NO 78

;; LENGTH: 802
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-245-730-78

Query Match 2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 82 KLVKNAQ 88
|||||
DB 413 KLVKNAQ 419

```
RESULT 25
US-10-245-739-78
; Sequence 78, Application US/10245739
; Publication No. US20030073195A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C86
; CURRENT APPLICATION NUMBER: US/10/245,739
; CURRENT FILING DATE: 2002-09-15
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-739-78

Query Match          2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      82 KLVKNAQ 88
        |||||||
Db      413 KLVKNAQ 419

RESULT 26
US-10-246-210-78
; Sequence 78, Application US/10246210
; Publication No. US20030073196A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
```

```
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C121
; CURRENT APPLICATION NUMBER: US/10/246,210
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-246-210-78

Query Match          2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      82 KLVKNAQ 88
        |||||||
Db      413 KLVKNAQ 419

RESULT 27
US-10-239-196-78
; Sequence 78, Application US/10239196
; Publication No. US20030074735A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C8
; CURRENT APPLICATION NUMBER: US/10/239,196
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
```



```

; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-239-196-78

Query Match
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 KLVKNAQ 88
      |||||
Db 413 KLVKNAQ 419

RESULT 28
US-10-243-024-78
; Sequence 78, Application US/10243024
; Publication No. US20030077741A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C38
; CURRENT APPLICATION NUMBER: US/10/243,024
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-024-78

Query Match
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-024-78

Query Match
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 KLVKNAQ 88
      |||||
Db 413 KLVKNAQ 419

RESULT 29
US-10-243-409-78
; Sequence 78, Application US/10243409
; Publication No. US20030077742A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C41
; CURRENT APPLICATION NUMBER: US/10/243,409
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-243-409-78

Query Match
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 KLVKNAQ 88
      |||||
```

Db 413 KLVKNAQ 419

```

RESULT 30
US-10-245-033-78
; Sequence 78, Application US/10245033
; Publication No. US20030078401A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephen, Jean-Phillippe
; APPLICANT: Matambe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C84
; CURRENT APPLICATION NUMBER: US/10/245,033
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-033-78

Query Match      2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 82 KLVKNAQ 88
Db 413 KLVKNAQ 419

RESULT 31
US-10-245-621-78
; Sequence 78, Application US/10245621
; Publication No. US20030077743A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin

```

```

; APPLICANT: Smith, Victoria
; APPLICANT: Stephen, Jean-Phillippe
; APPLICANT: Matambe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C104
; CURRENT APPLICATION NUMBER: US/10/245,621
; CURRENT FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-621-78

Query Match      2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 82 KLVKNAQ 88
Db 413 KLVKNAQ 419

RESULT 32
US-10-245-880-78
; Sequence 78, Application US/10245880
; Publication No. US20030077744A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephen, Jean-Phillippe
; APPLICANT: Matambe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C69
; CURRENT APPLICATION NUMBER: US/10/245,880
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114

```

```
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/063046
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/065027
;; PRIOR FILING DATE: 1997-11-10
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/086478
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090689
;; Remaining Prior Application data removed - See file wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 116
;; SEQ ID NO 78
;; LENGTH: 802
;; TYPE: prt
;; ORGANISM: Homo Sapien
US-10-245-880-78
```

```
Query Match
Best Local Similarity 2.4%; Score 7; DB 9; Length 802;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 82 KLVKNAQ 88
Db 413 KLVKNAQ 419
```

```
RESULT 33
US-10-243-095-78
;; Sequence 78, Application US/10243095
;; Publication No. US20030082726a1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin
;; APPLICANT: Eaton, Dan
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Phillippe
;; APPLICANT: Watanabe, Colin
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; APPLICANT: Fong, Sherman
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3630R1C43
;; CURRENT FILING DATE: 2002-09-12
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 10/197942
;; PRIOR FILING DATE: 2002-07-18
;; PRIOR APPLICATION NUMBER: 60/059114
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/063046
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/065027
;; PRIOR FILING DATE: 1997-11-10
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/086478
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090557
```

```
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090689
;; PRIOR FILING DATE: 1998-06-25
;; Remaining Prior Application data removed - See file wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 116
;; SEQ ID NO 78
;; LENGTH: 802
;; TYPE: prt
;; ORGANISM: Homo Sapien
US-10-243-095-78
```

```
Query Match
Best Local Similarity 2.4%; Score 7; DB 9; Length 802;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 82 KLVKNAQ 88
Db 413 KLVKNAQ 419
```

```
RESULT 34
US-10-245-185-78
;; Sequence 78, Application US/10245185
;; Publication No. US20030082728a1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin
;; APPLICANT: Eaton, Dan
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Phillippe
;; APPLICANT: Watanabe, Colin
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; APPLICANT: Fong, Sherman
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3630R1C80
;; CURRENT FILING DATE: 2002-09-16
;; PRIOR FILING DATE: 2002-07-18
;; PRIOR APPLICATION NUMBER: 10/197942
;; PRIOR FILING DATE: 2002-07-18
;; PRIOR APPLICATION NUMBER: 60/059114
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/063046
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/065027
;; PRIOR FILING DATE: 1997-11-10
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/086478
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090689
;; Remaining Prior Application data removed - See file wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 116
;; SEQ ID NO 78
;; LENGTH: 802
;; TYPE: prt
;; ORGANISM: Homo Sapien
US-10-245-185-78
```

```
Query Match
Best Local Similarity 2.4%; Score 7; DB 9; Length 802;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      82 KLVKNAQ 88
      |||||
Db      413 KLVKNAQ 419

RESULT 35
US-10-245-427-78
; Sequence 78, Application US/10245427
; Publication No. US20030082729A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C68
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78.
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-427-78

Query Match      2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      82 KLVKNAQ 88
      |||||
Db      413 KLVKNAQ 419

RESULT 36
US-10-245-473-78
; Sequence 78, Application US/10245473
; Publication No. US20030082730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
US-10-245-473-78

RESULT 37
US-10-245-770-78
; Sequence 78, Application US/10245770
; Publication No. US20030082731A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C74
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 78
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-473-78

Query Match      2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      82 KLVKNAQ 88
      |||||
Db      413 KLVKNAQ 419
```

PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 78
LENGTH: 802
TYPE: PRT
ORGANISM: Homo Sapien
US-10-245-770-78

Query Match 2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 KLVKNAQ 88
Db 413 KLVKNAQ 419

RESULT 38
US-10-245-877-78
Sequence 78, Application US/10245877
Publication No. US20030082732A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Watanabe, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3630R1C94
CURRENT APPLICATION NUMBER: US/10/245, 877
CURRENT FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 78
LENGTH: 802
TYPE: PRT
ORGANISM: Homo Sapien
US-10-245-877-78

Query Match 2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 KLVKNAQ 88
Db 413 KLVKNAQ 419

RESULT 39
US-10-246-976-78
Sequence 78, Application US/10246976
Publication No. US20030082732A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Watanabe, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3630R1C122
CURRENT APPLICATION NUMBER: US/10/246, 976
CURRENT FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 116
SEQ ID NO 78
LENGTH: 802
TYPE: PRT
ORGANISM: Homo Sapien
US-10-246-976-78

Query Match 2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 82 KLVKNAQ 88
|||||
DB 413 KLVKNAQ 419

RESULT 40
US-10-243-320-78

; Sequence 78, Application US/10243320
; Publication No. US20030087388A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin

; APPLICANT: Baton, Dan

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Goddard, Audrey

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin

; APPLICANT: Smith, Victoria

; APPLICANT: Stephen, Jean-Philippe

; APPLICANT: Matanbe, Colin

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zhenli

; APPLICANT: Fong, Sherman

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3630R1C48

; CURRENT APPLICATION NUMBER: US/10/243,320

; PRIOR FILING DATE: 2002-09-13

; PRIOR APPLICATION NUMBER: 10/197942

; PRIOR FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/059114

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/063046

; PRIOR FILING DATE: 1997-10-24

; PRIOR APPLICATION NUMBER: 60/065027

; PRIOR FILING DATE: 1997-11-10

; PRIOR APPLICATION NUMBER: 60/079689

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/086478

; PRIOR FILING DATE: 1998-05-22

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/089801

; PRIOR FILING DATE: 1998-06-18

; PRIOR APPLICATION NUMBER: 60/090557

; PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090689

; PRIOR FILING DATE: 1998-06-25

; Remaining Prior Application data removed - See file wrapper or PALM.

; NUMBER OF SEQ ID NOS: 116

; SEQ ID NO 78

; LENGTH: 802

; TYPE: PRT

; ORGANISM: Homo sapien

US-10-243-320-78

Query Match 2.4%; Score 7; DB 9; Length 802;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 82 KLVKNAQ 88
|||||
DB 413 KLVKNAQ 419

RESULT 41

US-09-853-253-5

; Sequence 5, Application US/09853253

; Patent No. US20020055156A1

; GENERAL INFORMATION:

; APPLICANT: JASPERS, STEPHEN
; APPLICANT: SHEPPARD, PAUL
; APPLICANT: DEISHER, THERESA

; APPLICANT: BISHOP, PAUL

; TITLE OF INVENTION: Zs1933-1like Peptides

; FILE REFERENCE: 00-30

; CURRENT APPLICATION NUMBER: US/09/853,253

; PRIOR FILING DATE: 2001-05-10

; PRIOR APPLICATION NUMBER: 60/203,300

; PRIOR FILING DATE: 2000-05-11

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 5

; LENGTH: 23

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-853-253-5

Query Match 2.1%; Score 6; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. .89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 73 EDGGQA 78
|||||
DB 9 EDGGQA 14

RESULT 42

US-09-853-253-6

; Sequence 6, Application US/09853253

; Patent No. US20020055156A1

; GENERAL INFORMATION:

; APPLICANT: JASPERS, STEPHEN

; APPLICANT: SHEPPARD, PAUL

; APPLICANT: DEISHER, THERESA

; APPLICANT: BISHOP, PAUL

; TITLE OF INVENTION: Zs1933-1like Peptides

; FILE REFERENCE: 00-30

; CURRENT APPLICATION NUMBER: US/09/853,253

; PRIOR FILING DATE: 2001-05-10

; PRIOR APPLICATION NUMBER: 60/203,300

; PRIOR FILING DATE: 2000-05-11

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 6

; LENGTH: 23

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: AMIDATION

; LOCATION: (23)...(23)

US-09-853-253-6

Query Match 2.1%; Score 6; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. .89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 73 EDGGQA 78
|||||
DB 9 EDGGQA 14

RESULT 43

US-09-853-253-4

; Sequence 4, Application US/09853253

; Patent No. US20020055156A1

; GENERAL INFORMATION:

; APPLICANT: JASPERS, STEPHEN

; APPLICANT: SHEPPARD, PAUL

; APPLICANT: DEISHER, THERESA

; APPLICANT: BISHOP, PAUL

; TITLE OF INVENTION: Zs1933-1like Peptides

; FILE REFERENCE: 00-30

```

; CURRENT APPLICATION NUMBER: US/09/853,253
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/203,300
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-853-253-4

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 EDGGOA 78
Db 9 EDGGOA 14

RESULT 44
US-10-140-164-34
; Sequence 34, Application US/10140164
; Publication No. US20030072736A1
; GENERAL INFORMATION:
; APPLICANT: Baker et al.
; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor TR16
; FILE REFERENCE: PPS1AC1
; CURRENT APPLICATION NUMBER: US/10/140,164
; CURRENT FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 09/637,856
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/148,348
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/148,683
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 60/148,870
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 60/148,758
; PRIOR FILING DATE: 1999-08-16
; PRIOR APPLICATION NUMBER: 60/149,181
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149,453
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/149,498
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 34
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-140-164-34

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GIYYSK 132
Db 5 GIYYSK 10

RESULT 45
US-10-140-164-63
; Sequence 63, Application US/10140164
; Publication No. US20030072736A1
; GENERAL INFORMATION:
; APPLICANT: Baker et al.
; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor TR16
; FILE REFERENCE: PPS1AC1
; CURRENT APPLICATION NUMBER: US/10/140,164
```

```

; CURRENT FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 09/637,856
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/148,348
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/148,683
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 60/148,870
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 60/148,758
; PRIOR FILING DATE: 1999-08-16
; PRIOR APPLICATION NUMBER: 60/149,181
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149,453
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/149,498
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 63
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-140-164-63

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GIYYSK 132
Db 5 GIYYSK 10
```

Search completed: May 21, 2003, 11:14:26
Job time : 27 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 10:57:11 ; Search time 11 Seconds

(without alignments)
1082.155 Million cell updates/sec

Title: US-09-869-677A-2

Perfect score: 287

Sequence: 1 SSTGAKTAKSDKLKAVATNS.....PGDSYAMKNNLKRISGL 287

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	231	80.5	310	1	M7SA_STRPY
2	36	11.5	309	1	M7SA_STRAP
3	36	11.5	309	1	M7SA_STRAP
4	33	11.5	309	1	M7SA_STRAP
5	28	9.8	310	1	M7A2_STRPN
6	27	9.4	309	1	M7A1_STRPN
7	27	9.4	309	1	M7SA_STRMT
8	20	7.0	309	1	M7SA_STRSA
9	20	7.0	310	1	M7SA_STRCR
10	20	7.0	310	1	M7SA_STRGC
11	19	6.6	313	1	M7SA_LACLA
12	16	5.6	306	1	M7SA_STRMT
13	9	3.1	310	1	M7A1_LISTIN
14	9	3.1	310	1	M7A1_LISTMO
15	8	2.8	383	1	NRL1_RHORH
16	8	2.8	586	1	DEGR_BACSU
17	7	2.4	60	1	DEGR_BACSU
18	7	2.4	139	1	YMYD_CAREL
19	7	2.4	143	1	ADX_CHICK
20	7	2.4	201	1	TDB_ECOLI
21	7	2.4	226	1	BASP_BOVIN
22	7	2.4	226	1	BASP_HUMAN
23	7	2.4	254	1	TIRI_YEAST
24	7	2.4	361	1	REL_COXBU
25	7	2.4	416	1	PGK_MYCGE
26	7	2.4	452	1	TIG_CAVCR
27	7	2.4	499	1	CPN1_RAT
28	7	2.4	500	1	CPN2_MESAU
29	7	2.4	500	1	CPN2_MOUSE
30	7	2.4	500	1	CPN3_MOUSE
31	7	2.4	510	1	CPN2_RAT
32	7	2.4	520	1	HPAB_ECOLI
33	7	2.4	520	1	HPAB_KLEPN

34	7	2.4	577	1	SYD_HELPY	P56459 helicobacte
35	7	2.4	599	1	ABE1_HUMAN	O96B10 homo sapien
36	7	2.4	600	1	V719_METUA	O58120 methanococc
37	7	2.4	659	1	POL_CERY	P05400 carnation e
38	7	2.4	674	1	POL_CAVWD	P03556 cauliflower
39	7	2.4	679	1	POL_CAVWC	P03555 cauliflower
40	7	2.4	679	1	POL_CAVWE	O02964 cauliflower
41	7	2.4	679	1	POL_CAVVS	P03554 cauliflower
42	7	2.4	680	1	POL_CAVVN	O00962 cauliflower
43	7	2.4	685	1	VARI_METUA	O57675 methanococc
44	7	2.4	862	1	ADHE_CLOAB	P33744 Clostridium
45	7	2.4	1134	1	YML7_YEAST	O03735 saccharomyc

ALIGNMENTS

RESULT 1	ID	MTSA_STRPY	STANDARD:	PRT:	310 AA.
AC	09A157	09RNU0	09RNT7		
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DE	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Metal ABC transporter substrate-binding lipoprotein precursor.				
GN	MTSA OR SPY0453.				
OS	Streptococcus pyogenes.				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_TaxID=1314;				
RN	[1]				
RP	SEQUENCE FROM N.A., SEQUENCE OF 30-39, AND CHARACTERIZATION.				
RC	STRAIN-SF370 / ATCC 700294 / Serotype M1;				
RX	MEDLINE-21192684; PubMed-11296296;				
RA	Ferretil J.J., McMan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,				
RA	Primeaux C., Szate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,				
RA	Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,				
RA	Xuan X., Clifton S.W., Roe B.A., McLaughlin R.,				
RT	*Complete genome sequence of an M1 strain of Streptococcus pyogenes.*;				
RL	Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).				
CC	-1- FUNCTION: Part of an ATP-driven transport system for a metal; this				
CC	protein has affinity for Zn(II), Fe(II) and Cu(II).				
CC	-1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID				
CC	ANCHOR.				
CC	-1- SIMILARITY: Belongs to the bacterial solute-binding protein family				
CC	9. Lipoprotein receptor antigen (LraI) subfamily.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
DR	EMBL: AF180520; AAD56936.1; -				
DR	EMBL: AF180521; AAD56939.1; -				
DR	EMBL: AE006505; AAK33468.1; -				
DR	HSSP: P96116; ITOA.				
DR	InterPro: IPR001987; Lipoprotein_4.				
DR	Pfam: PF01297; SBP_bac_9; 1.				
DR	PRINTS: PR00690; ADHESNFAMILY.				
KW	transport; zinc transport; iron transport; Copper; Membrane;				
KW	lipoprotein; signal; Complete proteome.				
FT	SIGNAL 1 20 PROBABLE.				

```

FT CHAIN 21 310 METAL ABC TRANSPORTER SUBSTRATE-
FT LIPID 21 21 BINDING LIPOPROTEIN.
FT VARIANT 77 21 N-ACYL DIGLYCERIDE (PROBABLE).
FT CONFLICT 26 26 T -> A (IN STRAIN AD1).
FT CONFLICT 30 30 K -> E (IN REF. 1).
FT CONFLICT 44 40 A -> G (IN REF. 1).
FT CONFLICT 49 50 A -> VM (IN REF. 1).
SQ SEQUENCE 310 AA; 34358 MW; B0F829EF1C72CADC CRC64;

Query Match 80.5%; Score 231; DB 1; Length 310;
Best Local Similarity 100.0%; Pred. No. 2e-228;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 EKSNADVIFRYNGINLEDGQAMFTKLKYNKQKTKNNDYFVADGIDVITLEGASKEGK 116
DB 78 EKSNADVIFRYNGINLEDGQAMFTKLKYNKQKTKNNDYFVADGIDVITLEGASKEGK 137
QY 117 DPHAMLENGIITSKNIATKOLAKOPKNETEKNKAYVAKLEKIDKEAKSKFDALAE 176
DB 138 DPHAMLENGIITSKNIATKOLAKOPKNETEKNKAYVAKLEKIDKEAKSKFDALAE 197
QY 177 NKRLIYSEGCFFYFSKAYGVPASAYIWEITEEGTPTDQSSLIETKLYIKPSALFVSS 236
DB 198 NKRLIYSEGCFFYFSKAYGVPASAYIWEITEEGTPTDQSSLIETKLYIKPSALFVSS 257
QY 237 VDRRPMEVYSKDSGIPYISIFPTDSIAKKGKPEDSYAMKRNLDKISBGL 287
DB 258 VDRRPMEVYSKDSGIPYISIFPTDSIAKKGKPEDSYAMKRNLDKISBGL 308

RESULT 2
MTSA_STRAP
ID MTSA_STRAP STANDARD; PRT; 309 AA.
AC 09LSW9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Manganese ABC transporter substrate-binding lipoprotein precursor
DE (Pneumococcal surface adhesin A).
GN PSAA.
OS Streptococcus anginosus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1328;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NCIC 10713;
RA MEDLINE=21418906; PubMed=11527799;
RA Jado I., Fenoll A., Casal J., Perez A.;
RT "Identification of the psaa gene, coding for pneumococcal surface
RT adhesin A, in viridans group streptococci other than Streptococcus
RT pneumoniae."
RL Clin. Diagn. Lab. Immunol. 8:895-898(2001).
CC -1- FUNCTION: Part of an ATP-driven transport system for manganese.
CC Also act as an adhesin which is involved on adherence to
CC extracellular matrix. It is an important factor in pathogenesis
CC and infection (By similarity).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
CC ANCHOR (By similarity).
CC -1- SIMILARITY: Belongs to the bacterial solute-binding protein family
CC 9. Lipoprotein receptor antigen (LraI) subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF248235; AAF64228.1;
CC HSSP: P96116; ITOA.
DR
SQ

```

```

DR InterPro: IP001987; Lipoprotein_4.
DR Pfam: PF01297; SBP_Dac_9; 1.
DR PRINTS: PR00690; ADHESNFAMILY.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN_1.
KW Transport; Manganese; Membrane; Lipoprotein; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 309 MANGANESE ABC TRANSPORTER SUBSTRATE-
FT LIPID 20 20 BINDING LIPOPROTEIN.
SQ SEQUENCE 309 AA; 34615 MW; 966EBDEB08F0580D CRC64;

Query Match 12.5%; Score 36; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 5.6e-29;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 KRLIYSEGCFFYFSKAYGVPASAYIWEITEEGTPT 213
DB 198 KRLIYSEGCFFYFSKAYGVPASAYIWEITEEGTPT 233

RESULT 3
MTSA_STROR
ID MTSA_STROR STANDARD; PRT; 309 AA.
AC 09LSW9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Manganese ABC transporter substrate-binding lipoprotein precursor
DE (Pneumococcal surface adhesin A).
GN PSAA.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NCIC 11427;
RA Jado I., Fenoll A., Casal J., Perez A.;
RT "Identification of the psaa gene, coding for pneumococcal surface
RT adhesin A, in viridans group streptococci other than Streptococcus
RT pneumoniae."
RL Clin. Diagn. Lab. Immunol. 8:895-898(2001).
CC -1- FUNCTION: Part of an ATP-driven transport system for manganese.
CC Also act as an adhesin which is involved on adherence to
CC extracellular matrix. It is an important factor in pathogenesis
CC and infection. It may contribute to the formation and accumulation
CC of dental plaque (By similarity).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
CC ANCHOR (By similarity).
CC -1- SIMILARITY: Belongs to the bacterial solute-binding protein family
CC 9. Lipoprotein receptor antigen (LraI) subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF248237; AAF64230.1;
CC HSSP: P96116; ITOA.
DR
DR InterPro: IP001987; Lipoprotein_4.
DR Pfam: PF01297; SBP_Dac_9; 1.
DR PRINTS: PR00690; ADHESNFAMILY.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN_1.
KW Transport; Manganese; Membrane; Lipoprotein; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 309 MANGANESE ABC TRANSPORTER SUBSTRATE-
FT LIPID 20 20 BINDING LIPOPROTEIN.
SQ SEQUENCE 309 AA; 34616 MW; 81F0C41F91E5C954 CRC64;

```

Query Match 12.5%; Score 36; DB 1; Length 309;
 Best Local Similarity 100.0%; Pred. No. 5,6e-29;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 178 KKLIVTSECFYFSKAYGVPSAYIWEINTEEGTP 213
 |||||
 DB 198 KKLIVTSECFYFSKAYGVPSAYIWEINTEEGTP 233

RESULT 4

MTSA_STRPA STANDARD; PRT; 309 AA.
 ID MTSA_STRPA
 AC P31305;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Manganese ABC transporter substrate-binding lipoprotein precursor
 DE (Adhesin B) (Saliva-binding protein).
 GN FIMA.
 OS Streptococcus parasanguis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OK NCBI_TaxID=1318;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FW213;
 RX MEDLINE=90035427; PubMed=2572555;
 RA Ferno J.C., Leblanc D.D., Fives-Taylor P.M.;
 RT "Nucleotide sequence analysis of a type 1 fimbrial gene of
 RT Streptococcus sanguis FW213."
 RL Infect. Immun. 57:3527-3533(1989).
 RN [2]
 RP POSSIBLE FUNCTION.
 RX MEDLINE=91147187; PubMed=1671775;
 RA Ganesbharan N., Hannam P.M., Kolenbrander P.E., McBride B.C.;
 RT "Nucleotide sequence of a gene coding for a saliva-binding protein
 RT (SAB) from Streptococcus sanguis 12 and possible role of the protein
 RT in coaggregation with actinomyces."
 RL Infect. Immun. 59:1093-1099(1991).
 CC -1- FUNCTION: Part of an ATP-driven transport system for manganese.
 CC Also act as an adhesin which is involved on adherence to
 CC extracellular matrix. It is an important factor in pathogenesis
 CC and infection. It may contribute to the formation and accumulation
 CC of dental plaque (by similarity).
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
 CC ANCHOR (By similarity).
 CC -1- SIMILARITY: Belongs to the bacterial solute-binding protein family
 CC 9. Lipoprotein receptor antigen (Lrai) subfamily.
 CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE A FIMBRIAL SUBUNIT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; M26130; AAS53077.1; -;
 DR PIR; A37186; A37186.
 DR HSSP; P96116; ITQA.
 DR InterPro; IPR001987; Lipoprotein_4.
 DR Pfam; PF01297; SBP_bac_9; 1.
 DR PRINTS; PR00690; ADHESNFAMILY.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Transport; Manganese; Membrane; Lipoprotein; Signal.
 FT SIGNAL 1 20 PROBABLE.
 FT CHAIN 21 309 MANGANESE ABC TRANSPORTER SUBSTRATE-
 FT BINDING LIPOPROTEIN.
 FT LIPID 21 21 N-ACYL DIGLYCERIDE (PROBABLE).
 FT SEQUENCE 309 AA; 34349 MW; 332F10BC858E8396 CRC64;

Query Match 11.5%; Score 33; DB 1; Length 309;
 Best Local Similarity 100.0%; Pred. No. 6,5e-26;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 181 IYTSSECFYFSKAYGVPSAYIWEINTEEGTP 213
 |||||
 DB 201 IYTSSECFYFSKAYGVPSAYIWEINTEEGTP 233

RESULT 5

MTA2_STRPN STANDARD; PRT; 310 AA.
 ID MTA2_STRPN
 AC P42363;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Manganese ABC transporter substrate-binding lipoprotein precursor
 DE (Pneumococcal surface adhesin A).
 GN PSA OR PAPA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OK NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-R36A;
 RX MEDLINE=94086122; PubMed=7505262;
 RA Sampson J.S., O'Connor S.P., Stinson A.R., Tharpe J.A., Russell H.;
 RT "Cloning and nucleotide sequence analysis of psa, the Streptococcus
 RT pneumoniae gene encoding a 37-kilodalton protein homologous to
 RT previously reported Streptococcus sp. adhesins."
 RL Infect. Immun. 62:319-324(1994).
 CC -1- FUNCTION: Part of an ATP-driven transport system for manganese.
 CC Also act as an adhesin which is involved on adherence to
 CC extracellular matrix. It is an important factor in pathogenesis
 CC and infection.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
 CC ANCHOR (By similarity).
 CC -1- SIMILARITY: Belongs to the bacterial solute-binding protein family
 CC 9. Lipoprotein receptor antigen (Lrai) subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; L19055; AAA16798.1; -;
 DR InterPro; IPR001987; Lipoprotein_4.
 DR Pfam; PF01297; SBP_bac_9; 1.
 DR PRINTS; PR00690; ADHESNFAMILY.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Transport; Manganese; Membrane; Lipoprotein; Signal.
 FT SIGNAL 1 20 PROBABLE.
 FT CHAIN 21 310 MANGANESE ABC TRANSPORTER SUBSTRATE-
 FT BINDING LIPOPROTEIN.
 FT LIPID 21 21 N-ACYL DIGLYCERIDE (PROBABLE).
 FT SEQUENCE 310 AA; 34538 MW; 9BBC8E84E572F8B8 CRC64;

Query Match 9.8%; Score 28; DB 1; Length 310;
 Best Local Similarity 100.0%; Pred. No. 8,5e-21;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 181 IYTSSECFYFSKAYGVPSAYIWEINTE 208
 |||||
 DB 202 IYTSSECFYFSKAYGVPSAYIWEINTE 229

RESULT 6
 ID MTA2_STRPN STANDARD; PRT; 309 AA.

AC P12538; Q54720; Q9R6P5; Q9L5X2; Q9L5X4; Q9L5X3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Manganese ABC transporter substrate-binding lipoprotein precursor
 DE (pneumococcal surface adhesin A).
 GN PSAA OR SP1650.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GB;
 RA Sampson J.S., Whitney A.M., Furrow Z.;
 RT "Streptococcus pneumoniae surface adhesin A.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RC STRAIN-D39 / NCTC 7466 / Serotype 2;
 RX MEDLINE=97101047; PubMed=8945574;
 RA Berry A.M., Paton J.C.;
 RT "Sequence heterogeneity of Psaa, a 37-kilodalton putative adhesin
 RT essential for virulence of Streptococcus pneumoniae.";
 RL Infect. Immun. 64:5255-5262(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98449534; PubMed=9767595;
 RA Novak R., Braun J.S., Charpentier E., Tuomanen E.;
 RT "Penicillin tolerance genes of Streptococcus pneumoniae: the ABC-type
 RT manganese peroxidase complex Psa.";
 RL Mol. Microbiol. 29:1285-1296(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NA-1064/97, NA-1283/96, NA-1383/97, NA-1508/92, and R6;
 RA Perez A., Jado I., Casal J.;
 RT "Identification of a psa gene in viridans streptococcal strains.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TIGR4;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Rattner D.,
 RA Holtzapple E., Khouri H., Wolf A.M., Ustebek T.R., Hansen C.L.,
 RA McDonald L.A., Feldlyum T.V., Angluoli S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae.";
 RL Science 293:498-506(2001).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=98025470; PubMed=9379902;
 RA Dintilhac A., Allong G., Granadel C., Claverys J.-P.;
 RT "Competence and virulence of Streptococcus pneumoniae: Adc and Psaa
 RT mutants exhibit a requirement for Zn and Mn resulting from
 RT inactivation of putative ABC metal permeases.";
 RL Mol. Microbiol. 25:727-739(1997).
 CC -1- FUNCTION: Part of an ATP-driven transport system for manganese.
 CC Also act as an adhesin which is involved on adherence to
 CC extracellular matrix. It is an important factor in pathogenesis
 CC and infection.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
 CC ANCHOR (By similarity).
 CC -1- SIMILARITY: Belongs to the bacterial solute-binding protein family
 CC 9. Lipoprotein receptor antigen (LraI) subfamily.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U53509; AAC09440.1; -
 DR EMBL: U40786; AAC24470.1; -
 DR EMBL: AF055088; AAD09975.1; -
 DR EMBL: AF248229; AAF70663.1; -
 DR EMBL: AF248230; AAF70664.1; -
 DR EMBL: AF248231; AAF70665.1; -
 DR EMBL: AF248232; AAF70666.1; -
 DR EMBL: AF248233; AAF70667.1; -
 DR EMBL: AF248234; AAF70668.1; -
 DR EMBL: AE007458; AAK75729.1; -
 DR HSSP: P96116; ITOA.
 DR TIGR: SP1650; -
 DR InterPro: IPR001987; Lipoprotein_4.
 DR Pfam: PF01297; SBP_bac_9; 1.
 DR PRINTS: PR00690; ADHESINFAMILY.
 DR PROSITE: PS00013; PROKR_LIPOPROTEIN; 1.
 DR TransProteome; Manganese; Membrane; Lipoprotein; Signal;
 KW Complete proteome.
 FT SIGNAL 1 19
 FT CHAIN 20 309
 FT
 FT LIPID 20 20
 FT
 FT VARIANT 8 8
 FT
 FT VARIANT 9 9
 FT
 FT VARIANT 14 14
 FT
 FT VARIANT 16 16
 FT
 FT VARIANT 16 16
 FT
 FT VARIANT 16 16
 FT
 FT VARIANT 27 28
 FT
 FT VARIANT 30 30
 FT
 FT VARIANT 62 62
 FT
 FT VARIANT 81 81
 FT
 FT VARIANT 83 83
 FT
 FT VARIANT 120 120
 FT
 FT VARIANT 130 130
 FT
 FT VARIANT 148 148
 FT
 FT VARIANT 164 164
 FT
 FT VARIANT 187 189
 FT
 FT VARIANT 193 193
 FT
 FT VARIANT 207 207
 FT
 FT VARIANT 234 234
 FT
 FT VARIANT 248 248
 FT
 FT VARIANT 285 285
 FT
 FT VARIANT 294 294
 FT
 FT SEQUENCE 309 AA; 34594 MW; B15E7EE3DA657C CRC64;
 SQ
 OY 29 IAGDKIDHSIPIQDDPHEIPLPED 55
 Db 49 IAGDKIDHSIPIQDDPHEIPLPED 75
 Query Match 9.4%; Score 27; DB 1; Length 309;
 Best Local Similarity 100.0%; Pred. No. 8,9e-20;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 7
 MTS_A_STRMT STANDARD; PRT; 309 AA.
 AC Q9L5X0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Manganese ABC transporter substrate-binding lipoprotein precursor

```

DE (Pneumococcal surface adhesin A).
GN PSMA.
OS Streptococcus mitis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=28037;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NCTC 12261;
RX MEDLINE-21418906; PubMed-11527799;
RA Jado I., Penoll A., Casal J., Perez A.;
RT "Identification of the psaA gene, coding for pneumococcal surface
RT adhesin A, in viridans group streptococci other than Streptococcus
RT pneumoniae."
RL Clin. Diagn. Immunol. 8:895-898(2001).
CC -1- FUNCTION: Part of an ATP-driven transport system for manganese.
CC Also act as an adhesin which is involved on adherence to
CC extracellular matrix. It is an important factor in pathogenesis
CC and infection (By similarity).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
CC ANCHOR (By similarity).
CC -1- SIMILARITY: Belongs to the bacterial solute-binding protein family
CC 9. Lipoprotein receptor antigen (Lrai) subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF248236; AAF64229.1; -.
DR HSSP; P96116; ITOA.
DR InterPro; IPR001987; Lipoprotein_4.
DR Pfam; PF01297; SBP_bac_9; 1.
DR PRINTS; PR00690; ADHESNFAMILY.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Transport; Manganese; Membrane; Lipoprotein; Signal.
FT SIGNAL 1 19 PROBABLE.
FT CHAIN 20 309 MANGANESE ABC TRANSPORTER SUBSTRATE-
FT BINDING LIPOPROTEIN.
FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
SQ SEQUENCE 309 AA; 34595 MW; 112423C9F4873D25 CRC64;

Query Match 9.4%; Score 27; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.9e-20;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 IAGDKIDLSIVPIGDPHEYPEPLPED 55
Db 49 IAGDKIDLSIVPIGDPHEYPEPLPED 75

RESULT 8
MTSA_STRSA STANDARD; PRT; 309 AA.
AC P31304;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Manganese ABC transporter substrate-binding lipoprotein precursor
DE (Adhesin B) (Saliva-binding protein).
GN SSAB.
OS Streptococcus sanguis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=1305;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 25-53.
RC STRAIN-12;
RX MEDLINE-9147187; PubMed-1671775;
RA Ganeshkumar N., Hannam P.M., Kolenbrander P.E., McBride B.C.;

```

```

RT "Nucleotide sequence of a gene coding for a saliva-binding protein
RT (Ssab) from Streptococcus sanguis 12 and possible role of the protein
RT in coaggregation with actinomyces."
RL Infect. Immun. 59:1093-1099(1991).
RN [2]
RP MUTAGENESIS.
RX MEDLINE-93123181; PubMed-8419308;
RA Ganeshkumar N., Arora N., Kolenbrander P.E.;
RT "Saliva-binding protein (Ssab) from Streptococcus sanguis 12 is a
RT lipoprotein."
RL J. Bacteriol. 175:572-574(1993).
CC -1- FUNCTION: Part of an ATP-driven transport system for manganese.
CC Also act as an adhesin which is involved on adherence to
CC extracellular matrix. It is an important factor in the
CC pathogenesis and infection. It may contribute to the formation and
CC accumulation of dental plaque.
CC -1- SUBUNIT: HOMODIMER AND HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
CC ANCHOR.
CC -1- SIMILARITY: Belongs to the bacterial solute-binding protein family
CC 9. Lipoprotein receptor antigen (Lrai) subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63481; AAC98426.1; -.
DR PIR; A43583; A43583.
DR HSSP; P96116; ITOA.
DR InterPro; IPR001987; Lipoprotein_4.
DR Pfam; PF01297; SBP_bac_9; 1.
DR PRINTS; PR00690; ADHESNFAMILY.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Transport; Manganese; Membrane; Lipoprotein; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 309 MANGANESE ABC TRANSPORTER SUBSTRATE-
FT BINDING LIPOPROTEIN.
FT LIPID 20 20 N-ACYL DIGLYCERIDE.
FT MUTAGEN 15 15 C->G; NO LOSS OF ACYLATION.
FT MUTAGEN 20 20 C->G; LOSS OF ACYLATION.
SQ SEQUENCE 309 AA; 34684 MW; 8FC8AF434AB18977 CRC64;

Query Match 7.0%; Score 20; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.3e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 SEKGKEDPHAWNLNGLIY 130
Db 131 SEKGKEDPHAWNLNGLIY 150

RESULT 9
MTSA_STRCR STANDARD; PRT; 310 AA.
AC Q53891;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metal ABC transporter substrate-binding lipoprotein precursor.
GN SCBA.
OS Streptococcus cristatus.
OC Plasmid pSCBA002.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=45634;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CC5A;
RX MEDLINE-96239025; PubMed-8675315;

```

RA Correla F.F., Dirienzo J.M., McKay T.L., Rosan B.;
 RT scda from Streptococcus cristata C53A: an atypical member of the lrai
 RT gene family.";
 RL Infect. Immun. 64:2114-2121(1996).
 CC -1- FUNCTION: Part of an ATP-driven transport system for manganese.
 CC Does not exhibit adhesion properties.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL SOLUTE-BINDING PROTEIN FAMILY
 CC 9.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: U46542; AAC44133.1; -;
 DR InterPro: IPR001987; Lipoprotein_4.
 DR Pfam: PF01297; SBP_bac_9; 1.
 DR PRINTS: PR00690; ADHESNFAMILY.
 DR PROSITE: PS00013; PROKAR.LIPOPROTEIN; 1.
 KM Transport: Manganese; Membrane; Lipoprotein; Signal; Plasmid.
 FT SIGNAL 1 19 PROBABLE.
 FT CHAIN 20 310 METAL ABC TRANSPORTER SUBSTRATE-BINDING
 FT LIPOPROTEIN.
 FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
 FT SEQUENCE 310 AA; 34725 MW; 14DBF7283ECF5936 CRC64;
 SQ
 Query Match 7.08; Score 20; DB 1; Length 310;
 Best Local Similarity 100.0%; Pred. No. 1.3e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 111 SEKGKEDPHAMLNENGIIT 130
 Db 132 SEKGKEDPHAMLNENGIIT 151
 RESULT 10
 MESA_STRGC STANDARD; PRT; 310 AA.
 AC PA2354;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metal ABC transporter substrate-binding lipoprotein precursor
 DE (Coaggregation-mediated adhesion).
 GN SCGA.
 OS Streptococcus gordonii Challis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CC NCB1_TaxID=29390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PK488;
 RX MEDLINE=95012638; PubMed=7927711;
 RA Kolenbrander P.E., Andersen R.N., Ganeshkumar N.;
 RT "Nucleotide sequence of the Streptococcus gordonii PK488
 RT coaggregation adhesion gene, scga, and ATP-binding cassette.";
 RL Infect. Immun. 62:4469-4480(1994).
 CC -1- FUNCTION: Part of an ATP-driven transport system for a metal;
 CC probably for manganese. Also act as an adhesin which is involved
 CC on adherence to extracellular matrix. It is an important factor in
 CC pathogenesis and infection.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
 CC ANCHOR (By similarity).
 CC -1- SIMILARITY: Belongs to the bacterial solute-binding protein family
 CC 9. Lipoprotein receptor antigen (lrai) subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: AED06363; AK05414.1; -;
 DR InterPro: IPR001987; Lipoprotein_4.
 DR Pfam: PF01297; SBP_bac_9; 1.
 DR PRINTS: PR00690; ADHESNFAMILY.
 DR PROSITE: PS00013; PROKAR.LIPOPROTEIN; 1.
 KM Transport: Manganese; Membrane; Lipoprotein; Signal;
 FT SIGNAL 1 23 PROBABLE.
 FT CHAIN 24 313 METAL ABC TRANSPORTER SUBSTRATE-
 FT BINDING LIPOPROTEIN.

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: L11577; AAA71947.1; -;
 DR InterPro: IPR001987; Lipoprotein_4.
 DR Pfam: PF01297; SBP_bac_9; 1.
 DR PRINTS: PR00690; ADHESNFAMILY.
 DR PROSITE: PS00013; PROKAR.LIPOPROTEIN; 1.
 KM Transport: Manganese; Membrane; Lipoprotein; Signal.
 FT SIGNAL 1 19 PROBABLE.
 FT CHAIN 20 310 METAL ABC TRANSPORTER SUBSTRATE-BINDING
 FT LIPOPROTEIN.
 FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
 FT SEQUENCE 310 AA; 34787 MW; D1DEB1A060BC252 CRC64;
 SQ
 Query Match 7.08; Score 20; DB 1; Length 310;
 Best Local Similarity 100.0%; Pred. No. 1.3e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 197 VPSAYWEINTEEGTPDOI 216
 Db 218 VPSAYWEINTEEGTPDOI 237
 RESULT 11
 MESA_LACLA STANDARD; PRT; 313 AA.
 AC Q9CFZ5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metal ABC transporter substrate-binding lipoprotein precursor.
 GN MESA OR LL1316.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 CC NCB1_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235166; PubMed=11337471;
 RA Bolotin A., Wincker P., Manger S., Jallion O., Malame K.,
 RA Weissenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis IL1403.";
 RL Genome Res. 11:731-753(2001).
 CC -1- FUNCTION: Part of an ATP-driven transport system for a metal;
 CC probably for manganese.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
 CC ANCHOR (By similarity).
 CC -1- SIMILARITY: Belongs to the bacterial solute-binding protein family
 CC 9. Lipoprotein receptor antigen (lrai) subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: AED06363; AK05414.1; -;
 DR InterPro: IPR001987; Lipoprotein_4.
 DR Pfam: PF01297; SBP_bac_9; 1.
 DR PRINTS: PR00690; ADHESNFAMILY.
 DR PROSITE: PS00013; PROKAR.LIPOPROTEIN; 1.
 KM Transport: Manganese; Membrane; Lipoprotein; Signal;
 FT SIGNAL 1 23 PROBABLE.
 FT CHAIN 24 313 METAL ABC TRANSPORTER SUBSTRATE-
 FT BINDING LIPOPROTEIN.

```

SQ      LIPID      24      24      N-ACYL DIGLYCERIDE (PROBABLE) .
SEQUENCE 313 AA; 35499 MW; 99146BF815B3BD30 CXC64;

Query Match
Best Local Similarity 100.0%; Score 19; DB 1; Length 313;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      198      PSAYIWEINTEEEGTPDOI 216
Db      222      PSAYIWEINTEEEGTPDOI 240

RESULT 12
ID      MTSA_STRMU      STANDARD;      PRT;      306 AA.
AC      09K13;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Metal ABC transporter substrate-binding lipoprotein precursor.
GN      SLOC.
OS      Streptococcus mutans.
OC      Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC      Streptococcus.
OX      NCBI_TaxID=1309;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 25175;
RX      MEDLINE=20359334; PubMed=10899841;
RA      Kitten T., Munro C.L., Michalek S.M., Macrina F.L.;
RT      "Genetic characterization of a Streptococcus mutans Lrai family
RL      operon and role in virulence."
RL      Infect. Immun. 68:4441-4451(2000).
CC      -1 FUNCTION: Part of an ATP-driven transport system for a metal;
CC      probably for manganese. Also act as an adhesin which is involved
CC      on adherence to extracellular matrix. It is an important factor in
CC      pathogenesis and infection. It may contribute to the formation and
CC      accumulation of dental plaque.
CC      -1 SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
CC      ANCHOR (By similarity).
CC      -1 SIMILARITY: Belongs to the bacterial solute-binding protein family
CC      9. Lipoprotein receptor antigen (Lrai) subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; AF232688; AAF81674.1; -
DR      InterPro; IPR001987; Lipoprotein_4.
DR      Pfam; PF01297; SBP_bac_9; 1.
DR      PRINTS; PR00690; ADHESNFM1Y.
DR      PROSITE; PS00013; PROKAR_LIPOROREIN_1.
KM      Transport; Manganese; Membrane; Lipoprotein; Signal.
FT      SIGNAL      1      19
FT      CHAIN      20      306      METAL ABC TRANSPORTER SUBSTRATE-
FT      LIPID      20      20      BINDING LIPOPROTEIN.
FT      SEQUENCE 306 AA; 34336 MW; 87B92E990DA644 CXC64;

Query Match
Best Local Similarity 100.0%; Score 16; DB 1; Length 306;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      198      PSAYIWEINTEEEGTP 213
Db      215      PSAYIWEINTEEEGTP 230

RESULT 13

```

ID	NTMA_LISTIN	STANDARD:	PRT:	310 AA.
AC	092AG1:			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Manganese-binding lipoprotein mntA precursor.			
GN	MNTA OR LIN1961.			
OS	Listeria innocua.			
OC	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.			
OX	NCBI_TaxID=1642;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CLIP 11262 / Serovar 66;			
RX	MEDLINE=21537279; PubMed=11679669;			
RA	Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,			
RA	Baquerot F., Berche P., Blocher H., Brandt P., Chakraborty T.,			
RA	Charbit A., Checourat S., Couve E., de Daruvar A., Deboux P.,			
RA	Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,			
RA	Enlian K.-D., Fshl H., Garcia-del Portillo F., Garrido P.,			
RA	Gautier L., Goebel W., Gomez-Lopez N., Hahn T., Hauf J., Jackson D.,			
RA	Jones L.-M., Kaestli U., Kreft J., Kuhn M., Kunst F., Kutapkat G.,			
RA	Madeno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,			
RA	Nordstlek G., Novella S., de Pablos B., Perez-Diaz J.-C., Puccell R.,			
RA	Rommel B., Rose M., Schlueter T., Simoes N., Tierrez A.,			
RA	Vaqueiro-Boland J.-A., Voss H., Weiland J., Cossart P.;			
RT	"Comparative genomics of <i>Listeria</i> species."			
RL	Science 294:849-852(2001).			
CC	-1- FUNCTION: This protein is probably a component of a manganese			
CC	permease, a binding protein-dependent, ATP-driven transport			
CC	system (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor			
CC	(Probable).			
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL SOLUTE-BINDING PROTEIN FAMILY			
CC	9.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib.ch).			
CC	-----			
DR	EMBL; AL596170; CAC97191.1; -			
DR	Listlist; LIN01961; -			
DR	InterPro: IPR001987; Lipoprotein_4.			
DR	Pfam: PF01297; SBP_bac_9; 1.			
DR	PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.			
KW	Transport; Metal-binding; Lipoprotein; Membrane; Signal;			
KW	Complete proteome.			
FT	SIGNAL	1	18	POTENTIAL.
FT	CHAIN	19	310	MANGANESE-BINDING LIPOPROTEIN MNTA.
FT	LIPID	19	19	N-ACTL DIGLYCERIDE (PROBABLE).
SO	SEQUENCE	310 AA;	34536 MW;	D90733E0A461F74 CRC64;
QY	200 AYIWEINTE	208		
DB	221 AYIWEINTE	229		
QY	3.1%; Score 9; DB 1;			
QY	Length 310;			
QY	Best Local Similarity 100.0%; Pred. No. 0.23;			
QY	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
RESULT 14				
NTMA_LISTMO	STANDARD:	PRT;	310 AA.	
AC	08Y653;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Manganese-binding lipoprotein mntA precursor.			

GN NMFA OR LMO1847.
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacilliales; Listeriaceae; Listeria.
 NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BD-E / Serovar 1/2a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Baquero F., Frangeul L., Buchrieser C., Rusnok C., Amend A.,
 Baquero F., Berche P., Biocker H., Brandt P., Chakraborty T.,
 Chardit A., Chetouani F., Couve E., de Daruvar A., Deloux P.,
 Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
 Ertlan K.-D., Fahl H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
 Madueno E., Maltounem A., Mata Vicente J., Ng E., Nedjari H.,
 Norddick G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
 RT "Comparative genomics of Listeria species."
 RL Science 294:849-852(2001).
 CC -1- FUNCTION: This protein is probably a component of a manganese
 CC permease, a binding protein-dependent, ATP-driven transport
 CC system (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (Probable).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL SOLUTE-BINDING PROTEIN FAMILY
 CC 9.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL591981; CAC99925.1; -
 DR Listlist: LMO01847; -
 DR InterPro: IPR001987; Lipoprotein_4.
 DR Pfam: PF01297; SPP_bac_9; 1.
 DR PRINTS: PR00690; ADHESNFAMILY.
 DR PROSITE: PS00013; PROKR_LIPOPROTEIN; 1.
 KW Transport; Metal-binding; Lipoprotein; Membrane; Signal;
 KM Complete proteome.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 310 MANGANESE-BINDING LIPOPROTEIN NMFA.
 FT LIPID 19 19 N-ACYL DIGLYCERIDE (PROBABLE).
 SQ SEQUENCE 310 AA: 34417 MW: 41899348A68C804 CRC64;
 Query Match 3.1%; Score 9; DB 1; Length 310;
 Best Local Similarity 100.0%; Pred. No. 0.23; Mismatches 0; Gaps 0;
 Matches 9; Conservative 0; Indels 0;
 Y 200 AYIWEINTE 208
 Db 221 AYIWEINTE 229
 Db 221 AYIWEINTE 229
 RESULT 15
 NR_L1_RHORN STANDARD; PRT; 383 AA.
 AC 002068;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Aliphatic nitrilase (EC 3.5.5.7).
 OS Rhodococcus rhodocrous.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Nocardiaceae; Rhodococcus.
 NCBI_TaxID=1829;
 RN [1]
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND MUTAGENESIS OF CYS-170.

RC STRAIN-K22;
 RX MEDLINE=93003039; PubMed=1390687;
 RA Kobayashi M., Yanaka N., Nagasawa T., Yamada H.;
 RT "Primary structure of an aliphatic nitrile-degrading enzyme,
 RT aliphatic nitrilase, from Rhodococcus rhodocrous K22 and expression
 RT of its gene and identification of its active site residue."
 RL Biochemistry 31:9000-9007(1992).
 CC -1- FUNCTION: ACTS ON ALIPHATIC NITRILES SUCH AS ACRYLONITRILE,
 CC CROTONONITRILE AND GLUTARONITRILE.
 CC -1- CATALYTIC ACTIVITY: R-CN + H(2)O = R-COOH + NH(3).
 CC -1- SIMILARITY: BELONGS TO THE NITRILASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D12583; BAA02127.1; -
 DR PIR: A43470; A43470.
 DR InterPro: IPR000132; N/CN_hydrolase.
 DR InterPro: IPR003010; Ntlase/CNhydrlse.
 DR Pfam: PF00795; CN_hydrolase; 1.
 DR PROSITE: PS00920; NITRIL_CHT_1; 1.
 DR PROSITE: PS00921; NITRIL_CHT_2; 1.
 KW Hydrolase.
 FT ACT_SITE 170 170 PROBABLE.
 FT MUTAGEN 170 170 C->S/A: 100% LOSS OF ACTIVITY.
 SQ SEQUENCE 383 AA: 42275 MW: BA8E5572B8DA17B CRC64;
 Query Match 2.8%; Score 8; DB 1; Length 383;
 Best Local Similarity 100.0%; Pred. No. 2.9; Mismatches 0; Gaps 0;
 Matches 8; Conservative 0; Indels 0;
 Y 50 EPLPEDAE 57
 Db 272 EPLPEDAE 279
 Db 272 EPLPEDAE 279
 RESULT 16
 HOL1_YEAST STANDARD; PRT; 586 AA.
 ID HOL1_YEAST
 AC P53389;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE HOL1 protein;
 GN HOL1 OR YNR035C OR N3494.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97113542; PubMed=8955402;
 RA Wright M.B., Howell E.A., Gaber R.F.;
 RT "Amino acid substitutions in membrane-spanning domains of Hol1, a
 RT member of the major facilitator superfamily of transporters, confer
 RT nonselective cation uptake in Saccharomyces cerevisiae."
 RL J. Bacteriol. 178:7197-7205(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Pohl T.M.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-305 FROM N.A.
 RA Duesterhoeft A., Floeth M., Fritz C., Heuss-Meltzel D.,
 RA Hilbert H., Moestl D.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: SEEMS TO BE INVOLVED IN THE UPTAKE OF SEVERAL CATIONS
 CC AND OF HISTIDINOL.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
 CC AS THE DRUG RESISTANCE TRANSDUCASE FAMILY). DHA14 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: I42348; AAB47713.1; -
 DR EMBL: 271670; CAA96336.1; -
 DR EMBL: 271671; CAA96338.1; -
 DR SGD: S0005338; HOL1.
 DR Transport; Transmembrane.
 KW TRANSMEM 67
 FT TRANSMEM 104 124 POTENTIAL.
 FT TRANSMEM 131 151 POTENTIAL.
 FT TRANSMEM 190 210 POTENTIAL.
 FT TRANSMEM 220 240 POTENTIAL.
 FT TRANSMEM 363 383 POTENTIAL.
 FT TRANSMEM 414 434 POTENTIAL.
 FT TRANSMEM 449 469 POTENTIAL.
 FT TRANSMEM 478 498 POTENTIAL.
 FT TRANSMEM 509 529 POTENTIAL.
 FT TRANSMEM 545 565 POTENTIAL.
 FT TRANSMEM 510 510 POTENTIAL.
 FT CONFLICT 510 L -> F (IN REF. 1).
 SQ SEQUENCE 586 AA; 65348 MW; BDB6415256C3600F CRC64;
 Query Match 2.8%; Score 8; DB 1; Length 586;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 123 NLENGIIV 130
 DB 24 NLENGIIV 31
 RESULT 17
 DEGR_BACSU STANDARD; PRT; 60 AA.
 AC P06563;
 GN P06563;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Regulatory protein degr.
 DE DEGR OR PTRR.
 OS Bacillus subtilis, and
 OS Bacillus subtilis var. natto.
 OC Bacteria: Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423; 86029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-B.subtilis;
 RX MEDLINE=87083406; PubMed=3098734;
 RA Yang M., Shimotsu H., Ferrari E., Henner D.J.;
 RT "Characterization and mapping of the Bacillus subtilis ptrr gene";
 RL J. Bacteriol. 169:434-437(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-B.subtilis; STRAIN=168 / Marburg;
 RX MEDLINE=96349105; PubMed=8760912;
 RA Sorokin A.V., Azevedo V., Zmistein E., Galleron N., Ehrlich S.D.,
 RA Seror P.;
 RT "Sequence analysis of the Bacillus subtilis chromosome region between
 RT the sera and kdg loci cloned in a yeast artificial chromosome";
 RL Microbiology 142:2005-2016(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-B.s.natto;
 RX MEDLINE=86168015; PubMed=3082853;

RA Nagami Y., Tanaka T.;
 RT "Molecular cloning and nucleotide sequence of a DNA fragment from
 RT Bacillus natto that enhances production of extracellular proteases
 RT and levanucrase in Bacillus subtilis";
 RL J. Bacteriol. 166:20-28(1986).
 CC -1- FUNCTION: LEADS, IN B.SUBTILIS, TO ENHANCED PRODUCTION OF
 CC LEVANUCLEASE, ALKALINE PROTEASE, AND NEUTRAL PROTEASE. IN B.NATTO
 CC IT IS NONESSENTIAL FOR GROWTH OR EXPRESSION OF PROTEASES AND
 CC LEVANUCLEASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M15318; AAA22672.1; -
 DR EMBL: L77246; AAA96622.1; -
 DR EMBL: M12917; AAA22671.1; -
 DR EMBL: 299115; CAB14112.1; -
 DR Subtilist; Bg10699; degr.
 KW Transcription regulation; Complete proteome.
 SQ SEQUENCE 60 AA; 7109 MW; C7B7F8922F0AB131 CRC64;
 Query Match 2.4%; Score 7; DB 1; Length 60;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 262 IAKGKP 268
 DB 27 IAKGKP 33
 RESULT 18
 YMYD_CAEEL STANDARD; PRT; 139 AA.
 AC 09BIB4;
 GN 09BIB4;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Transcription initiation factor IYA small chain homolog.
 GN B0336.13.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Tach A.;
 RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE TFIIA-GAMMA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U32305; AAK1861.1; -
 DR HSP; P32774; 1YTF.
 DR Wormpep; B0336.13; CE26429.
 DR InterPro; IPR003194; TFIIA_gamma.
 DR Pfam; PF02268; TFIIA_gamma; 1.

DR Pfam: PF02751; TFIIF, gamma C, 1.
 KW Hypothetical protein; Transcription regulation; Nuclear protein.
 SQ SEQUENCE 139 AA; 15570 MW; 7D468A1E80F895D CRC64;
 Query Match 2.4%; Score 7; DB 1; Length 139;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 DKIKVVA 17
 |||||
 Db 95 DKIKVVA 101

RESULT 19
 ADX_CHICK
 ID ADX_CHICK STANDARD; PRT; 143 AA.
 AC P13216;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Adrenodoxin, mitochondrial precursor (adrenal ferredoxin) (Fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 ON NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=88326329; PubMed=3415692;
 RA Kagimoto K., McCarthy J.L., Waterman M.R., Kagimoto M.;
 RT "Purified amino acid sequence of mature chicken testis ferredoxin.";
 RL Biochem. Biophys. Res. Commun. 155:379-383(1988).
 RN [2]
 RP SEQUENCE OF 4-143 FROM N.A.
 RC TISSUE=Kidney;
 RA Blanchard R.K., Henry H.L.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP STRUCTURE BY NMR.
 RX MEDLINE=91369915; PubMed=1909889;
 RA Stjefeldt L., Markley J.L., Coghlan V.M., Vickery L.E.;
 RT "1H NMR spectra of vertebrate [2Fe-2S] ferredoxins. Hyperfine
 resonances suggest different electron delocalization patterns from
 plant ferredoxins.";
 RL Biochemistry 30:9078-9083(1991).
 CC -1- FUNCTION: ADRENODOXIN TRANSFERS ELECTRONS FROM ADRENODOXIN
 REDUCTASE TO THE CHOLESTEROL SIDE CHAIN CLEAVAGE CYTOCHROME P450.
 CC -1- COFACTOR: BINDS 1 2FE-2S CLUSTER.
 CC -1- SIMILARITY: BELONGS TO THE ADRENODOXIN / PUTIDAREDOXIN FAMILY.
 CC -1- SIMILARITY: BELONGS TO THE ADRENODOXIN / PUTIDAREDOXIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: M21275; AAA8576.1; -;
 DR EMBL: U25823; AAA82597.1; -;
 DR PIR: A31574; A31574.
 DR HSSP: P00257; 1E6E.
 DR InterPro: IPR001055; Adrenodoxin.
 DR InterPro: IPR001041; Ferredoxin.
 DR Pfam: PF00111; fer2; 1.
 DR PROSITE: PS00814; ADX; 1.
 KW Iron-sulfur; Electron transport; Mitochondrion; Transit peptide.
 FT NON_TER 1 1
 FT TRANSIT <1 19 MITOCHONDRION.
 FT CHAIN 20 143 ADRENODOXIN.
 FT METAL 65 65 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

FT METAL 71 71 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 74 74 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT METAL 111 111 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
 FT CONFLICT 4 4 V -> P (IN REF. 2).
 FT CONFLICT 16 16 A -> S (IN REF. 2).
 SQ SEQUENCE 143 AA; 15482 MW; 15BC3DA3067914B7 CRC64;
 Query Match 2.4%; Score 7; DB 1; Length 143;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 265 GKPGDS 271
 |||||
 Db 41 GKPGDS 47

RESULT 20
 TTDB_ECOLI
 ID TTDB_ECOLI STANDARD; PRT; 201 AA.
 AC P05851; P33131;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE L(+)-tartrate dehydratase beta subunit (EC 4.2.1.32) (L-TTD beta).
 GN TTDB OR B3062.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 ON NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / W3110;
 RX MEDLINE=93381464; PubMed=8371115;
 RA Reaney S.K., Begg C., Bungard S.I., Guest J.R.;
 RT "Identification of the L-tartrate dehydratase genes (tda and tdb)
 of Escherichia coli and evolutionary relationship with the class I
 fumarate genes.";
 RL J. Gen. Microbiol. 139:1523-1530(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87248073; PubMed=3297921;
 RA Nesin M., Iupski J.R., Svec P., Godson G.N.;
 RT "Possible new genes as revealed by molecular analysis of a 5-kb
 Escherichia coli chromosomal region 5' to the rpsU-dnaG-rpoD
 macromolecular-synthesis operon.";
 RL Gene 51:149-161(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9276503;
 RA Blatter F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.D.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:123-147(1997).
 CC -1- CATALYTIC ACTIVITY: (R,R)-tartrate + H(2)O.
 CC -1- SUBUNIT: TRIMER OF TWO ALPHA AND TWO BETA SUBUNITS (PROBABLY).
 CC -1- SIMILARITY: TO THE C-TERMINAL OF FUMARASES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: L14781; AAA03062.1; -;
 DR EMBL: M16194; AAA72574.1; -;
 DR EMBL: U28379; AAA89142.1; -;
 DR EMBL: AE000388; AAC76098.1; -;
 DR PIR: B29049; QOECRZ.

DR Ecogene; EG11169; ttdb.
 DR InterPro; IPR004647; ttdb_funa_fumb.
 DR TIGRfams; TIGR00723; ttdb_funa_fumb; 1.
 KW Lyase; Complete proteome.
 FT ACT_SITE 37
 FT CONFLICT 80 85 POTENTIAL.
 FT CONFLICT 165 165 QYAGK -> SMRES (IN REF. 3).
 FT CONFLICT 165 165 A -> V (IN REF. 3).
 SQ SEQUENCE 201 AA; 22640 MW; 25C14DCD1913F3CE CRC64;
 Query Match 2.4%; Score 7; DB 1; Length 201;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 174 IAEKKL 180
 |||||
 Db 175 IAEKKL 181
 RESULT 21
 BASP_BOVIN STANDARD; PRT; 226 AA.
 AC P80724; 002692;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Brain acid soluble protein 1 (BASP1 protein) (Neuronal axonal membrane protein NAP-22).
 GN BASP1 OR NAP22.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
 OC NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Brain;
 RX MEDLINE-97455984; PubMed-9310187;
 RA Mosevitsky M.I., Capony J.-P., Skladchikova G.Y., Novitskaya V.A., Plekhanov A.Y., Zakharov V.V.;
 RT "The BASP1 family of myristoylated proteins abundant in axonal terminal. Primary structure analysis and physico-chemical properties.";
 RT Biochimie 79:373-384(1997).
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Park S., Kim B., Jeong H., Kim Y.-I., Lee S., Oh Y., Baek K., Yoon J.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE ANCHORED. ASSOCIATED WITH THE MEMBRANES OF "GROWTH CONES" THAT FORM THE TIPS OF ELONGATING AXONS.
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE BASP/NAP22 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL; U92535; AAC67307.1; -;
 KW Membrane; Myristate; Neurone.
 FT INIT_MET 0
 FT LIPID 1
 FT CONFLICT 113 114 MYRISTATE.
 FT CONFLICT 127 131 AA -> LR (IN REF. 2).
 FT CONFLICT 127 131 APAES -> GRRPR (IN REF. 2).
 SQ SEQUENCE 226 AA; 22879 MW; 43665E0D9A046DD CRC64;
 Query Match 2.4%; Score 7; DB 1; Length 226;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 207 TEEBGP 213
 |||||
 AC P10863;
 RESULT 23
 TIRL_YEAST STANDARD; PRT; 254 AA.
 ID TIRL_YEAST

Db 30 TEEBGP 36
 RESULT 22
 BASP_HUMAN STANDARD; PRT; 226 AA.
 ID P80723; 043596;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Brain acid soluble protein 1 (BASP1 protein) (Neuronal axonal membrane protein NAP-22) (Neuronal tissue-enriched acidic protein).
 GN BASP1 OR NAP22.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Brain;
 RX MEDLINE-97455984; PubMed-9310187;
 RA Mosevitsky M.I., Capony J.-P., Skladchikova G.Y.U., Novitskaya V.A., Plekhanov A.Y.U., Zakharov V.V.;
 RT "The BASP1 family of myristoylated proteins abundant in axonal terminal. Primary structure analysis and physico-chemical properties.";
 RT Biochimie 79:373-384(1997).
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Park S., Kim Y.-I., Kim B., Jeong C., Oh Y., Baek K., Yoon J.;
 RL "Characterization of bovine and human cDNAs encoding NAP-22 (22 kDa neuronal tissue-enriched acidic protein) homologs.";
 RL Mol. Cells 8:471-477(1998).
 CC -1- SUBCELLULAR LOCATION: MEMBRANE ANCHORED. ASSOCIATED WITH THE MEMBRANES OF "GROWTH CONES" THAT FORM THE TIPS OF ELONGATING AXONS.
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- MASS SPECTROMETRY: MW-22780; METHOD-Electrospray.
 CC -1- SIMILARITY: BELONGS TO THE BASP/NAP22 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL; AF039656; AAC67374.1; -;
 KW Membrane; Myristate; Neurone.
 FT INIT_MET 0
 FT LIPID 1
 FT CONFLICT 44 44 MYRISTATE.
 FT CONFLICT 113 114 AA -> P (IN REF. 2).
 FT CONFLICT 127 131 AA -> LR (IN REF. 2).
 FT CONFLICT 127 131 APAES -> GRRPR (IN REF. 2).
 SQ SEQUENCE 226 AA; 22562 MW; 51A290E07B4C32F8 CRC64;
 Query Match 2.4%; Score 7; DB 1; Length 226;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 207 TEEBGP 213
 |||||
 Db 30 TEEBGP 36
 RESULT 23
 TIRL_YEAST STANDARD; PRT; 254 AA.
 ID TIRL_YEAST

DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Cold shock induced protein TIR1 precursor (Serine-rich protein 1).
 GN TIR1 OR SRP1 OR YER11W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89011972; PubMed=3139887;
 RA Marguet D., Guo X.J., Lauglin G.J.-M.;
 RT "Yeast gene SRP1 (serine-rich protein). Intragenic repeat structure
 and identification of a family of SRP1-related DNA sequences.";
 RL J. Mol. Biol. 202:455-470(1988).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288c / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
 RA Hyman R., Kayser A., Kemp C., Lashkari D., Lew H., Lin D.,
 RA Moseedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
 RA Petzel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 [3]
 RP PRESENCE OF A PIR1/2/3 REPEAT.
 RA Bairoch A.;
 RL Unpublished observations (OCT-1993).
 [4]
 RP INDUCTION.
 RX MEDLINE=95264920; PubMed=7746155;
 RA Kowalski L.R.Z., Kondo K., Inouye M.;
 RT "Cold-shock induction of a family of TIR1-related proteins associated
 with the membrane in Saccharomyces cerevisiae.";
 RL Mol. Microbiol. 15:341-353(1995).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 (potential).
 CC -1- INDUCTION: BY COLD SHOCK AND GLUCOSE.
 CC -1- PTM: O-GLYCOSYLATED (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE SRP1 / TIR1 FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 PIR1/2/3 REPEAT.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X12775; CAA31262.1; -
 DR EMBL: U18778; AAB64544.1; -
 DR PIR: S05803; PUBS.
 DR SGD: S0000813; TIR1.
 DR InterPro: IPR000992; SRP1_TIR1.
 DR InterPro: IPR000420; Yeast_PIR.
 DR Pfam: PF00399; Yeast_PIR.1.
 DR Pfam: PF00660; SRP1_TIR1.1.
 DR PROSITE: PS00929; PIR_REPEAT_1.1.
 DR PROSITE: PS00256; PIR_REPEAT_2.1.
 DR PROSITE: PS00724; SRP1_TIR1.1.
 DR Repeat; Signal; Glycoprotein; GPI-anchor.
 KW Repeat; Signal; Glycoprotein; GPI-anchor.
 FT CHAIN 1 18 POTENTIAL.
 FT CHAIN 19 233 COLD SHOCK INDUCED PROTEIN TIR1.
 FT PROPEP 234 254 REMOVED IN MATURE FORM (POTENTIAL).
 FT DOMAIN 113 254 ALA/SER-RICH.
 FT REPEAT 210 224 PIR1/2/3 REPEAT.
 FT LIPID 233 233 GPI-ANCHOR (POTENTIAL).
 SQ SEQUENCE 254 AA; 24906 MW; 7BB8F8399758BDEE CRC64;

Query Match 2.4%; Score 7; DB 1; Length 254;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SSTGAKT 7
 Db 204 SSTGAKT 210
 RESULT 24
 ID RE1_COXBU STANDARD; PRT; 361 AA.
 AC P47849;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Peptide chain release factor 1 (RF-1).
 GN PREA.
 OS Coxiella burnetii.
 CC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
 CC Coxiella group; Coxiella.
 OX NCBI_TaxID=777;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nine Mile phase I;
 RX MEDLINE=96139650; PubMed=8549730;
 RA Williams H., Thiele D., Krauss H.;
 RT "Sequencing and linkage analysis of a Coxiella burnetii 2.1 kb NOT1
 fragment.";
 RL Eur. J. Epidemiol. 11:559-561(1995).
 CC -1- FUNCTION: PEPTIDE CHAIN RELEASE FACTOR 1 DIRECTS THE TERMINATION
 CC OF TRANSLATION IN RESPONSE TO THE PEPTIDE CHAIN TERMINATION
 CC CODONS UAG AND UAA (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC AND MITOCHONDRIAL RELEASE
 CC FACTORS FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X78969; CAA55563.1; -
 DR InterPro: IPR005139; PCRF.
 DR InterPro: IPR000352; PIR_rel_factor_1.
 DR InterPro: IPR004373; PIR1.
 DR Pfam: PF00472; RF-1; 1.
 DR Pfam: PF03462; PCRF; 1.
 DR TIGRPFAM: TIGR00019; PIR1.
 DR PROSITE: PS00745; RF_PROK_1.1.
 KW Protein biosynthesis.
 SQ SEQUENCE 361 AA; 40749 MW; 1BB0A26074088391 CRC64;
 QY 218 SLTEKLK 224
 Db 4 SLTEKLK 10
 RESULT 25
 ID PGK_MYCBE STANDARD; PRT; 416 AA.
 AC P47542; Q49338; Q59521;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphoglycerate kinase (EC 2.7.2.3).
 Query Match 2.4%; Score 7; DB 1; Length 361;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GN PKR OR MG300.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A. / G-37;
 RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischman R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fitchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrman J.L.,
 RA Tomb J.-F., Dougherty B.A., Boff K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium.";
 RL Science 270:397-403(1995).
 RN [2]
 RP SEQUENCE OF 1-55; 204-311 AND 314-414 FROM N.A.
 RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Boff K.F., Hutchison C.A. III;
 RT "A survey of the Mycoplasma genitalium genome by using random
 sequencing.";
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -1- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate -> ADP + 3-
 CC phospho-D-glyceroyl phosphate.
 CC -1- PATHWAY: Second phase of glycolysis; second step.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U39710; AAC1522.1; -;
 DR EMBL: U02226; AA03379.1; -;
 DR EMBL: U02234; AA03386.1; -;
 DR EMBL: U02178; AAD1246.1; ALF_INIT.
 DR HSSP: P36204; IYPE.
 DR TIGR: MG300; -;
 DR InterPro: IPR001576; PKG.
 DR Pfam: PF00167; PKG.1.
 DR PRINTS: PR00477; PHGLICKINASE.
 DR PROSITE: PS00111; GLYCERATE_KINASE; 1.
 KW Transferrase; Kinase; Glycolysis; Complete proteome.
 FT CONFLICT 204 206 SPO -> KPT (IN REF. 2).
 SQ SEQUENCE 416 AA; 45379 MW; 2A76E2C294BD2E6 CRC64;
 QY
 Db 10 SDKLKVV 16 2.4%; Score 7; DB 1; Length 416;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 219 SDKLKVV 225
 RESULT 26
 TIG-CAUCR STANDARD; PRT; 452 AA.
 AC 087705;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Trigger factor (TF).
 GN TIG OR CC1964.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;

OC Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A. / CB15;
 RC STRAIN-ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phade N.D., Ely B.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Knout H., Shetty J., Berry K.,
 RA Ueberberg T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 RN [2]
 RP SEQUENCE OF 119-452 FROM N.A.
 RC STRAIN-CB15N / NA1000;
 RA Jenal U., Fuchs T.;
 RL Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT. ACTS AS A CHAPERONE BY
 CC MAINTAINING THE NEWLY SYNTHESIZED PROTEIN IN AN OPEN CONFORMATION
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY. TIG SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE005869; AA23939.1; -;
 DR EMBL: A010321; CA09089.1; -;
 DR TIGR: CC1964; -;
 DR InterPro: IPR001179; FKBP_PPIASE.
 DR Pfam: PF00254; FKBP.1.
 DR TIGRFAMs: TIGR00115; tlg.1.
 DR PROSITE: PS00453; FKBP_PPIASE_1; FALSE_NEG.
 DR PROSITE: PS00454; FKBP_PPIASE_2; FALSE_NEG.
 DR PROSITE: PS50059; FKBP_PPIASE_3; 1.
 KW Cell division; Chaperone; Isomerase; Rotamase; Complete proteome.
 FT DOMAIN 169 254 PPIASE, FKBP-TYPE.
 FT CONFLICT 215 215 P -> R (IN REF. 2).
 SQ SEQUENCE 452 AA; 49767 MW; 5C05DB6A4B2DAD94 CRC64;
 QY
 Db 53 PEDAEKT 59 2.4%; Score 7; DB 1; Length 452;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 335 PEDAEKT 341
 RESULT 27
 CPNL-RAT STANDARD; PRT; 499 AA.
 AC P15393; Q64655;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytochrome P450 11B1, mitochondrial precursor (P450 11 beta-DS).
 GN CYP11B1 OR CYP11B-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-Sprague-Dawley;
 RX MEDLINE-90005967; PubMed-2551730;
 RA Nonaka Y., Matsukawa N., Morohashi K., Omura T., Ogihara T.,
 RT Teraoka H., Okamoto M.;
 RA "Molecular cloning and sequence analysis of cDNA encoding rat adrenal
 RT cytochrome P-450(11)beta."
 RL FEBS Lett. 255:21-26(1989).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Dahl salt-resistant; TISSUE-Adrenal gland;
 RX MEDLINE-93232078; PubMed-8473350;
 RA Matsukawa N., Nonaka Y., Higaki J., Naganu M., Mikami H., Ogihara T.,
 RT Okamoto M.;
 RA "Dahl's salt-resistant normotensive rat has mutations in cytochrome
 RT P450(11 beta), but the salt-sensitive hypertensive rat does not."
 RL J. Biol. Chem. 268:9117-9121(1993).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE-Testis;
 RX MEDLINE-93232080; PubMed-8473352;
 RA Mukai K., Imai M., Shimada H., Ishimura Y.;
 RT "Isolation and characterization of rat CYP11B genes involved in late
 RL steps of mineralo- and glucocorticoid syntheses."
 RL J. Biol. Chem. 268:9130-9137(1993).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Salt-resistant variant;
 RX MEDLINE-95352519; PubMed-7626522;
 RA Okamoto M., Nonaka Y., Ohta M., Takemori H., Halder S. K.,
 RT Zhi-Nong W., Sun T., Hatano O., Takakura A., Murakami T.;
 RT "Cytochrome P450(11 beta): structure-function relationship of the
 RL enzyme and its involvement in blood pressure regulation."
 RL J. Steroid Biochem. Mol. Biol. 53:89-94(1995).
 [5]
 RP SEQUENCE OF 25-44.
 RC TISSUE-Adrenal cortex;
 RX MEDLINE-89291826; PubMed-2738055;
 RA Ogihara T., Mitani F., Ishimura Y.;
 RT "Isolation of aldosterone synthase cytochrome P-450 from zona
 RL glomerulosa mitochondria of rat adrenal cortex."
 RL J. Biol. Chem. 264:10935-10938(1989).
 CC -1- FUNCTION: FORMS CORTICOSTERONE FROM 11-DEOXYCORTICOSTERONE.
 CC -1- CATALYTIC ACTIVITY: A steroid + reduced adrenal ferredoxin + O(2)
 CC = an 11-beta-hydroxysteroid + oxidized adrenal ferredoxin + H(2)O.
 CC -1- TISSUE SPECIFICITY: ADRENAL MITOCHONDRIAL.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR PROSITE: P500086; CYTOCHROME_P450; 1.
 KW Electron transport; Oxidoreductase; Monooxygenase; Membrane;
 KW Heme; Mitochondrion; Steroidogenesis; Transit peptide.
 FT TRANSIT 1 24
 FT CHAIN 25 499
 FT BINDING 446 446
 FT VARIANT 127 127
 FT VARIANT 351 351
 FT VARIANT 381 381
 FT VARIANT 384 384
 FT VARIANT 443 443
 SQ SEQUENCE 499 AA; 57458 MW; 19C041F96A673C7E CRC64;
 V -> M (IN SALT-RESISTANT STRAIN).
 V -> L (IN SALT-RESISTANT STRAIN).
 V -> C (IN SALT-RESISTANT STRAIN).
 R -> C (IN SALT-RESISTANT STRAIN).
 HEME (BY SIMILARITY).
 MITOCHONDRION.
 CYTOCHROME P450 11B1.

Query Match 2.4%; Score 7; DB 1; Length 499;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 52 LPEDAEK 58
 |||||
 DB 93 LPEDAEK 99

RESULT 28
 ID CPN2_MESAU STANDARD; PRT; 500 AA.
 AC 064658;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytochrome P450 11B2, mitochondrial precursor (RC 1.14.15.4)
 DE (C450X1B2) (Aldosterone synthase) (P-450A1do) (ALDOs) (Aldosterone-
 DE synthesizing enzyme) (Steroid 18-hydroxylase) (P-450C18).
 GN CYP11B2.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OX NCBI_TaxID=10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Adrenal glomerulosa;
 RX MEDLINE-94304774; PubMed-8031709;
 RA Lehoux J.-G., Mason J.I., Bernard H., Ducharme L., Lehoux J.,
 RA Veronneau S., Lefebvre A.;
 RT "The presence of two cytochrome P450 aldosterone synthase mRNAs in
 RT the hamster adrenal."
 RL J. Steroid Biochem. Mol. Biol. 49:131-137(1994).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97447701; PubMed-9303441;
 RA Coulombe N., Lefebvre A., Lehoux J.-G.;
 RT "Characterization of the hamster CYP11B2 gene encoding adrenal
 RT cytochrome P450 aldosterone synthase."
 RL DNA Cell Biol. 16:993-1002(1997).
 CC -1- FUNCTION: CONVERTS 11-DEOXYCORTICOSTERONE INTO CORTICOSTERONE,
 CC 18-HYDROXCORTICOSTERONE, AND ALDOSTERONE. ALSO CAN CATALYZE THE
 CC CONVERSION OF 11-DEOXYCORTISOL TO CORTISOL, 18-HYDROXYCORTISOL AND
 CC CORTISONE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: A steroid + reduced adrenal ferredoxin + O(2)
 CC = an 11-beta-hydroxysteroid + oxidized adrenal ferredoxin + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- TISSUE SPECIFICITY: ADRENAL GLAND.
 CC -1- INDUCTION: BY LOW SODIUM INTAKE.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: S73910; AAB31349.1; -
 DR EMBL: U71280; AAB16805.1; -
 DR HSSP: P00189; 1SCC.
 DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450.1.
 DR PRINTS: PR00385; P450.1.
 DR PROSITE: PS00086; CYTOCHROME_P450.1.
 KW Electron transport; Oxidoreductase; Monooxygenase; Membrane;
 Heme; Mitochondrion; Steroidogenesis; Transit peptide.
 FT TRANSIT 1 24
 FT CHAIN 25 500 CYTOCHROME P450 11B2.
 FT BINDING 447 500 HEME (BY SIMILARITY).
 SQ SEQUENCE 500 AA: 5731 MM; 3825IEDAD2085DE8 CRC64;

Query Match 2.4%; Score 7; DB 1; Length 500;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 52 LPEDAEK 58
 |||||
 DB 93 LPEDAEK 99

RESULT 29
 CPN2_MOUSE STANDARD; PRT; 500 AA.
 ID CPN2_MOUSE
 AC P15539; 064661;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytochrome P450 11B2, mitochondrial precursor (EC 1.14.15.4) (CYPX1B2)
 DE (P450C11) (Steroid 11-beta-hydroxylase) (Aldosterone synthase).
 GN CYP11B2 OR CYP11B-2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92168022; PubMed-1686470;
 RA Donalik L.J., Chaplin D.D., Kirkman M.S., Wu R.C., Liu W.,
 RA Howard T.A., Seldin M.F., Parker K.L.;
 RT "Different isozymes of mouse 11 beta-hydroxylase produce
 RT mineralocorticoids and glucocorticoids.";
 RL Mol. Endocrinol. 5:1853-1861(1991).
 RN [2]
 RP SEQUENCE OF 1-42 FROM N.A.
 RX MEDLINE-89093122; PubMed-2783417;
 RA Mouw A.R., Rice D.A., Meade J.C., Chua S.C., White P.C.,
 RA Schlimer B.P., Parker K.L.;
 RT "Structural and functional analysis of the promoter region of the
 RT gene encoding mouse steroid 11 beta-hydroxylase.";
 RL J. Biol. Chem. 264:1305-1309(1989).
 CC -1- FUNCTION: FORMS CORTICOSTERONE FROM 11-DEOXYCORTICOSTERONE.
 CC -1- CATALYTIC ACTIVITY: A steroid + reduced adrenal ferredoxin + O(2)
 CC "an 11-beta-hydroxysteroid + oxidized adrenal ferredoxin + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL: S85260; AAB21517.2; -
 DR EMBL: J04451; AAA50299.1; -
 DR PIR: A41552; A41552.
 DR HSSP: P00189; 1SCC.
 DR MGI: 88584; Cyp11b2.
 DR InterPro: IPR001128; Cytochrome_P450.

DR Pfam: PF00067; P450.1.
 DR PRINTS: PR00385; P450.1.
 DR PROSITE: PS00086; CYTOCHROME_P450.1.
 KW Electron transport; Oxidoreductase; Monooxygenase; Membrane;
 Heme; Mitochondrion; Steroidogenesis; Transit peptide.
 FT TRANSIT 1 24
 FT CHAIN 25 500 CYTOCHROME P450 11B2.
 FT BINDING 447 500 HEME (BY SIMILARITY).
 FT CONFLICT 366 366 K -> R (IN REF. 1; AAB21517).
 FT CONFLICT 383 383 G -> E (IN REF. 1; AAB21517).
 SQ SEQUENCE 500 AA: 57315 MM; 9381313CED5AB7A3 CRC64;

Query Match 2.4%; Score 7; DB 1; Length 500;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 52 LPEDAEK 58
 |||||
 DB 93 LPEDAEK 99

RESULT 30
 CPN3_RAT STANDARD; PRT; 500 AA.
 ID CPN3_RAT
 AC P30100;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytochrome P450 11B3, mitochondrial precursor (EC 1.14.15.4) (CYPX1B3)
 DE (P450-Aldo-2) (Aldosterone synthase).
 GN CYP11B3 OR CYP11B-3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Adrenal gland;
 RX MEDLINE-90274674; PubMed-2350348;
 RA Matukawa N., Nonaka Y., Ying Z., Higaki J., Ogihara T., Okamoto M.;
 RT "Molecular cloning and expression of cDNAs encoding rat aldosterone
 RT synthase: variants of cytochrome P-450(11 beta).";
 RL Biochem. Biophys. Res. Commun. 169:245-252(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-9222921; PubMed-1562515;
 RA Okamoto M., Nonaka Y.;
 RT "Molecular biology of rat steroid 11 beta-hydroxylase [P450(11 beta)]
 RT and aldosterone synthase [P450(11 beta)].";
 RL J. Steroid Biochem. Mol. Biol. 41:415-419(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Adrenal gland;
 RX MEDLINE-90242993; PubMed-2129527;
 RA Imai M., Shimada H., Okada Y., Matsushima-Hibiya Y., Ogishima T.,
 RA Ishimura Y.;
 RT "Molecular cloning of a cDNA encoding aldosterone synthase cytochrome
 RT P-450 in rat adrenal cortex.";
 RL FEBS Lett. 263:299-302(1990).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-9332080; PubMed-8473352;
 RA Mukai K., Imai M., Shimada H., Ishimura Y.;
 RT "Isolation and characterization of rat CYP11B genes involved in late
 RT steps of mineralo- and glucocorticoid syntheses.";
 RL J. Biol. Chem. 268:9130-9137(1993).
 RN [5]
 RP SEQUENCE OF 25-44.
 RC STRAIN-Sprague-Dawley; TISSUE-Adrenal gland;
 RX MEDLINE-89291826; PubMed-2738055;
 RA Ogishima T., Mitani F., Ishimura Y.;
 RT "Isolation of aldosterone synthase cytochrome P-450 from zona
 RT glomerulosa mitochondria of rat adrenal cortex.";


```

RL J. Biol. Chem. 264:10935-10938(1989).
CC -1- FUNCTION: CONVERTS 11-DEOXYCORTICOSTERONE INTO CORTICOSTERONE.
CC 18-HYDROXYCORTICOSTERONE. AND ALDOSTERONE. ALSO CAN CATALYZE THE
CC CONVERSION OF 11-DEOXYCORTISOL TO CORTISOL, 18-HYDROXYCORTISOL AND
CC CORTISONE.
CC -1- CATALYTIC ACTIVITY: A steroid + reduced adrenal ferredoxin + O(2)
CC = an 11-beta-hydroxysteroid + oxidized adrenal ferredoxin + H(2)O.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- TISSUE SPECIFICITY: ADRENAL CORTEX.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL: D00568; BAA00445.1; -
DR EMBL: X52766; CAA36978.1; -
DR EMBL: D14097; BAA03172.1; -
DR EMBL: D14097; BAA03172.1; JOINED.
DR EMBL: D14093; BAA03172.1; JOINED.
DR EMBL: D14094; BAA03172.1; JOINED.
DR EMBL: D14095; BAA03172.1; JOINED.
DR EMBL: D14096; BAA03172.1; JOINED.
DR PIR: A34281; A34281.
DR PIR: S09736; S09736.
DR PIR: B35342; B35342.
DR HSSP: P00189; ISCC.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450.1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450.1.
KW Electron transport; Oxidoreductase; Monooxygenase; Membrane;
KW Heme; Mitochondrion; Steroidogenesis; Transit peptide.
FT TRANSIT 1 24 MITOCHONDRION
FT CHAIN 25 500 CYTOCHROME P450 11B3.
FT BINDING 447 447 HEME (BY SIMILARITY).
FT CONFLICT 310 310 K -> E (IN REF. 3 AND 4).
SO SEQUENCE 500 AA; 57121 MW; 040F6CC8A4CDEAD CRC64;

Query Match 2.4%; Score 7; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 LPEDAER 58
DB 93 LPEDAER 99

RESULT 31
CPN2-RAT STANDARD; PRT; 510 AA.
AC P30099; 064540; (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome P450 11B2, mitochondrial precursor (EC 1.14.15.4) (CYPX1B2)
DE (P450-Aldo-1) (Aldosterone synthase).
GN CYP11B2 OR CYP11B-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE=90274674; PubMed=2350348;
RA Matsukawa N., Nonaka Y., Ying Z., Higaki J., Ogiwara T., Okamoto M.;
RT "Molecular cloning and expression of cDNAs encoding rat aldosterone
RT synthase: variants of cytochrome P-450(11 beta).";

```

```

RL Biochem. Biophys. Res. Commun. 169:245-252(1990).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=92222921; PubMed=1562515;
RA Okamoto M., Nonaka Y.;
RT "Molecular biology of rat steroid 11 beta-hydroxylase (P450(11 beta))
RT and aldosterone synthase (P450(11 beta, aldo)).";
RL J. Steroid Biochem. Mol. Biol. 41:415-419(1992).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Adrenal gland;
RX MEDLINE=93326104; PubMed=833830;
RA Zhou M., Gomez-Sanchez C.E.;
RT "Cloning and expression of a rat cytochrome P-450 11 beta-
RT hydroxylase/aldosterone synthase (CYP11B2) cDNA variant.";
RL Biochem. Biophys. Res. Commun. 194:112-117(1993).
RN [4]
RN SEQUENCE OF 35-54.
RC TISSUE=Adrenal cortex;
RX MEDLINE=89291826; PubMed=2738055;
RA Ogishima T., Mitani F., Ishimura Y.;
RT "Isolation of aldosterone synthase cytochrome P-450 from zona
RL glomerulosa mitochondria of rat adrenal cortex.";
CC J. Biol. Chem. 264:10935-10938(1989).
CC -1- FUNCTION: CONVERTS 11-DEOXYCORTICOSTERONE INTO CORTICOSTERONE.
CC 18-HYDROXYCORTICOSTERONE. AND ALDOSTERONE. ALSO CAN CATALYZE THE
CC CONVERSION OF 11-DEOXYCORTISOL TO CORTISOL, 18-HYDROXYCORTISOL AND
CC CORTISONE.
CC -1- CATALYTIC ACTIVITY: A steroid + reduced adrenal ferredoxin + O(2)
CC = an 11-beta-hydroxysteroid + oxidized adrenal ferredoxin + H(2)O.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- TISSUE SPECIFICITY: ADRENAL CORTEX.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL: D00567; BAA00444.1; -
DR EMBL: U14908; AAB60457.1; -
DR PIR: A34281; A34281.
DR PIR: A35342; A35342.
DR HSSP: P00189; ISCC.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450.1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450.1.
KW Electron transport; Oxidoreductase; Monooxygenase; Membrane;
KW Heme; Mitochondrion; Steroidogenesis; Transit peptide.
FT TRANSIT 1 34 MITOCHONDRION
FT CHAIN 35 510 CYTOCHROME P450 11B2.
FT BINDING 457 457 HEME (BY SIMILARITY).
FT VARIANT 146 146 E -> D.
FT VARIANT 261 261 Q -> R.
FT VARIANT 509 509 I -> V.
FT CONFLICT 1 13 MGACNDPFLHS -> MKKAPAKL (IN REF. 3).
SO SEQUENCE 510 AA; 58240 MW; 2E5129E513DEA9E CRC64;

Query Match 2.4%; Score 7; DB 1; Length 510;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 LPEDAER 58
DB 103 LPEDAER 109

RESULT 32
HPAB_ECOLI

```



```

ID  HPAB_ECOLI  STANDARD;  PRT;  520 AA.
AC  057160;
DT  15-JUN-2002 (Rel. 41, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last sequence update)
DE  4-hydroxyphenylacetate 3-monooxygenase, oxygenase component (EC
DE  1.14.14.-) (4-HPA 3-monooxygenase large component) (4-HPA 3-
DE  hydroxylase).
GN  HPAB.
OS  Escherichia coli.
OC  Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC  Escherichia.
OX  NCBI_TaxID=562;
RN  [1]
RP  SEQUENCE FROM N.A. AND SEQUENCE OF 1-20.
RC  STRAIN=M / ATCC 11105;
RX  MEDLINE=94357932; PubMed=8077235;
RA  Prieto M.A., Garcia J.L.;
RT  "Molecular characterization of 4-hydroxyphenylacetate 3-hydroxylase of
RT  Escherichia coli. A two-protein component enzyme.";
RL  J. Biol. Chem. 269:22823-22829(1994).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=W / ATCC 11105;
RX  MEDLINE=96135235; PubMed=8550403;
RA  Prieto M.A., Diaz E., Garcia J.L.;
RT  "Molecular characterization of the 4-hydroxyphenylacetate catabolic
RT  pathway of Escherichia coli W: engineering a mobile aromatic
RT  degradative cluster.";
RL  J. Bacteriol. 178:111-120(1996).
RN  [3]
RP  FUNCTION.
RX  MEDLINE=20120494; PubMed=10653707;
RA  Yun L., Sandvik E.R.;
RT  "Characterization of 4-hydroxyphenylacetate 3-hydroxylase (HpaB) of
RT  Escherichia coli as a reduced flavin adenine dinucleotide-utilizing
RT  monooxygenase.";
RL  Appl. Environ. Microbiol. 66:481-486(2000).
CC  -1- FUNCTION: Utilizes FADH(2) supplied by hpaC or by another flavin
CC  reductase, to catalyze the hydroxylation of 4-hydroxyphenylacetic
CC  acid, leading to the production of 3,4-DHPA. Can also oxidize
CC  phenol to catechol, and hydroxylate other phenol derivatives.
CC  -1- CATALYTIC ACTIVITY: 4-hydroxyphenylacetate + FAD + H(2)O = 3,4-
CC  dihydroxyphenylacetate + FAD + H(2)O.
CC  -1- PATHWAY: 4-hydroxyphenylacetate catabolism; first step.
CC  -1- SUBUNIT: 4-HPA 3-monooxygenase consists of a reductase component
CC  hpaC and an oxygenase component hpaB.
CC  -1- INDUCTION: By 4-hydroxyphenylacetic acid.
CC  -1- MISCELLANEOUS: E.coli K12 lacks the oxygenase component hpaB.
CC  -1- SIMILARITY: BELONGS TO THE FADH(2)-UTILIZING MONOOXYGENASE FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; Z37980; CAAB6048.1; -
DR  EMBL; Z29081; CAAB2321.1; -
DR  InterPro; IPR004925; HpaB.
KW  Aromatic hydrocarbons catabolism; Oxidoreductase; Monooxygenase; FAD.
SQ  SEQUENCE 520 AA; 58447 MW; 1E93BBC29BBB55C CRC64;

Query Match 2.4%; Score 7; DB 1; Length 520;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 14 KVVATNS 20
DB 192 KVVATNS 198

```

```

RESULT 33
HPAB_KIEPN
ID  HPAB_KIEPN  STANDARD;  PRT;  520 AA.
AC  048440;
DT  15-JUN-2002 (Rel. 41, Created)
DT  15-JUN-2002 (Rel. 41, Last sequence update)
DE  4-hydroxyphenylacetate 3-monooxygenase, oxygenase component (EC
DE  1.14.14.-) (4-HPA 3-monooxygenase large component) (4-HPA 3-
DE  hydroxylase).
GN  HPAB OR HPAB.
OS  Klebsiella pneumoniae.
OC  Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC  Klebsiella.
OX  NCBI_TaxID=573;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=M5a1;
RX  MEDLINE=97195731; PubMed=9133323;
RA  Gibello A., Suarez M., Allende J.L., Martin M.;
RT  "Molecular cloning and analysis of the genes encoding the 4-
RT  hydroxyphenylacetate hydroxylase from Klebsiella pneumoniae.";
RL  Arch. Microbiol. 167:160-166(1997).
CC  -1- FUNCTION: Utilizes FADH(2) supplied by hpaC or by another flavin
CC  reductase, to catalyze the hydroxylation of 4-hydroxyphenylacetic
CC  acid, leading to the production of 3,4-DHPA (By similarity).
CC  -1- CATALYTIC ACTIVITY: 4-hydroxyphenylacetate + FADH(2) + O(2) = 3,4-
CC  dihydroxyphenylacetate + FAD + H(2)O.
CC  -1- PATHWAY: 4-hydroxyphenylacetate catabolism; first step.
CC  -1- SUBUNIT: 4-HPA 3-monooxygenase consists of a reductase component
CC  hpaC and an oxygenase component hpaB (By similarity).
CC  -1- INDUCTION: By 3- or 4-hydroxyphenylacetic acid.
CC  -1- SIMILARITY: BELONGS TO THE FADH(2)-UTILIZING MONOOXYGENASE FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL; L41068; AAC37120.1; -
DR  Pfam; PF03241; HpaB. 1.
KW  Aromatic hydrocarbons catabolism; Oxidoreductase; Monooxygenase; FAD.
SQ  SEQUENCE 520 AA; 58651 MW; 0EA03932BA7121EA CRC64;

Query Match 2.4%; Score 7; DB 1; Length 520;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 14 KVVATNS 20
DB 192 KVVATNS 198

RESULT 34
STD_HELPY
ID  STD_HELPY  STANDARD;  PRT;  577 AA.
AC  P56459;
DT  15-JUL-1998 (Rel. 36, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DE  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE  (AspRS).
GN  ASPS OR HP0617.
OS  Helicobacter pylori (Campylobacter pylori).
OC  Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC  Helicobacter.
OX  NCBI_TaxID=210;
RN  [1]
RP  SEQUENCE FROM N.A.

```


CC endoribonuclease activity. May play a central role in the
 CC regulation of mRNA turnover. Antagonizes the anti-viral effect of
 CC the interferon-regulated 2-5A/RNase L pathway. May act as a
 CC chaperone for post-translational events during HIV-1 capsid
 CC assembly.
 CC -1 SUBUNIT: Probably heterodimerizes with RNase L. Associates with
 CC HIV-1 GAG, GAG-POL and VIF capsid proteins.
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial. Localized to
 CC clusters of virus formation at the plasma membrane.
 CC -1 INDUCTION: Activated by encephalomyocarditis virus (EMCV) and HIV-
 CC 1.
 CC -1 MISCELLANEOUS: The ABC transporter domains seem not to be
 CC functional.
 CC -1 SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. ABCG1
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: X76388; CAA53972.1; -;
 CC DR EMBL: X74987; CAA52920.1; -;
 CC DR EMBL: BC016283; AAH16283.1; -;
 CC DR EMBL: BC016988; AAH16988.1; -;
 CC DR EMBL: U90446; AAC24730.1; -;
 CC DR EMBL: BC005422; AAH05422.1; -;
 CC DR InterPro: IPR003439; ABC_transporter.
 CC Pfam: PF00005; ABC_tran; 2.
 CC ProDom: PD000006; ABC_transportr; 2.
 CC SMART: SM00382; AAA; 2.
 CC DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 CC DR MIM: 601213; -;
 CC DR MGD: MGI:1195458; Abce1.
 CC KM Mitochondrion; Chaperone; ATP-binding; Repeat.
 CC FT NP_BIND 110 117 ATP (POTENTIAL).
 CC FT NP_BIND 379 386 ATP (POTENTIAL).
 CC FT CONFLICT 118 118 T -> A (IN REF. 1).
 CC FT CONFLICT 153 153 Q -> K (IN REF. 3).
 CC FT CONFLICT 158 158 R -> T (IN REF. 3).
 CC FT CONFLICT 174 179 DOIPKA -> ARFLRL (IN REF. 1).
 CC FT CONFLICT 471 473 ALA -> RLR (IN REF. 1).
 CC FT SEQUENCE 599 AA; 67314 MW; 5D582B62E95BC7A6 CRC64;
 CC
 CC Query Match 2.4%; Score 7; DB 1; Length 599;
 CC Best Local Similarity 100.0%; Pred. No. 45;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 195 YGVPSAY 201
 CC Db 290 YGVPSAY 296
 CC
 CC RESULT 36
 CC ID V719_METUA STANDARD; PRT; 600 AA.
 CC AC 058129; METUA
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Hypothetical ABC transporter ATP-binding protein M0719.
 CC GN M0719.
 CC OS Methanococcus jannaschii.
 CC OC Archaeae; Euryarchaeota; Methanococci; Methanococcales;
 CC OC Methanocaldococcaceae; Methanocaldococcus.
 CC OX NCBI_TaxID=2190;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 CC MEDLINE=9633799; PubMed=8688087;

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.A., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weldman J.F., Fuhmann J.L., Nguyen D.,
 RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RA jannaschii";
 RA Science. 273:1058-1073(1996).
 CC
 CC -1 SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: U67518; AAB98713.1; -;
 CC DR TIGR: M0719; -;
 CC DR InterPro: IPR003593; AAA_Arpase.
 CC DR InterPro: IPR003439; ABC_transporter.
 CC Pfam: PF00005; ABC_tran; 2.
 CC Pfam: PF00037; fer4; 1.
 CC DR PRINTS: PR00353; 4FE4SFROXIN.
 CC DR PRODOM: PD000006; ABC_transportr; 2.
 CC DR SMART: SM00382; AAA; 2.
 CC DR PROSITE: PS00198; 4FE4S_FERREDOXIN; UNKNOWN_1.
 CC DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 CC KM Hypothetical protein; Transport; ATP-binding; Complete proteome.
 CC FT NP_BIND 117 124 ATP (POTENTIAL).
 CC FT NP_BIND 380 387 ATP (POTENTIAL).
 CC FT SEQUENCE 600 AA; 68018 MW; 07BC8DF6107B354C CRC64;
 CC
 CC Query Match 2.4%; Score 7; DB 1; Length 600;
 CC Best Local Similarity 100.0%; Pred. No. 45;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 195 YGVPSAY 201
 CC Db 293 YGVPSAY 299
 CC
 CC RESULT 37
 CC ID POL_CERV STANDARD; PRT; 659 AA.
 CC AC P05400; POL_CERV
 CC DT 01-NOV-1988 (Rel. 09, Created)
 CC DT 01-NOV-1988 (Rel. 09, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Enzymatic polypeptide [Contains: Aspartic protease (EC 3.4.23.-);
 CC DE Endonuclease; Reverse transcriptase (EC 2.7.7.49)].
 CC GN V.
 CC OS Carnation etched ring virus (CERV).
 CC OC Viruses; Retroviruses; Caulimoviridae; Caulimovirus.
 CC OX NCBI_TaxID=10640;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RA Hull R., Sadler J., Longstaff M.;
 CC RT "The sequence of carnation etched ring virus DNA: comparison with
 CC cauliflower mosaic virus and retroviruses.";
 CC EMBO J. 5:3083-3090(1986).
 CC
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + [DNA](N).
 CC -1- SIMILARITY: HIGH, WITH OTHER CAULIMOVIRUS ORF V.
 CC -1- SIMILARITY: WITH RETROVIRAL POL/GAG POLYPROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -


```

QY 223 LKVIKPS 229
    |||||
Db 272 LKVIKPS 278

RESULT 40
POL_CAMVE STANDARD; PRT; 679 AA.
AC 002964;
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enzymatic polypeptide [Contains: Aspartic protease (EC 3.4.23.-);
DE Endonuclease; Reverse transcriptase (EC 2.7.7.49)].
GN V.
OS Cauliflower mosaic virus (strain BRC) (CamV).
OC Viruses; Retroid viruses; Caulimoviridae; Caulimovirus.
OX NCBI_TaxID=31556;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93154593; PubMed=8428667;
RA Chenault K.D., Melcher U.K.;
RT "The complete nucleotide sequence of cauliflower mosaic virus isolate
RT BRC."
RL Gene 123:255-257(1993).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate
CC + [DNA](N).
CC -1- SIMILARITY: HIGH, WITH OTHER CAULIMOVIRUS ORF V.
CC -1- SIMILARITY: WITH RETROVIRAL POL/GAG POLYPEPTIDES.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M90542; AAA62375.1; -.
DR MEROPS: A03.001; -.
DR InterPro: IPR000588; Peptidase_A3.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; RVT; 1.
DR Pfam: PF02160; Peptidase_A3; 1.
DR PRINTS: PR00731; CAULIMOPASE.
KW Hydrolyase; Aspartyl protease; RNA-directed DNA polymerase;
KW Transferase; Endonuclease; Polypeptide.
FT ACT_SITE 45 45 BY SIMILARITY.
FT SIMILAR 40 130 TO RETROVIRUS GAG/POL PROTEASE DOMAIN.
FT SIMILAR 260 620 TO RETROVIRUS GAG/POL DNA POLYMERASE
FT DOMAIN.
SQ SEQUENCE 679 AA; 78713 MM; 2D8031B62F1CB05A CRC64;

Query Match 2.4%; Score 7; DB 1; Length 679;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 223 LKVIKPS 229
    |||||
Db 272 LKVIKPS 278

RESULT 41
POL_CAMVS STANDARD; PRT; 679 AA.
AC P03554;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enzymatic polypeptide [Contains: Aspartic protease (EC 3.4.23.-);
DE Endonuclease; Reverse transcriptase (EC 2.7.7.49)].
GN V.

```

```

OS Cauliflower mosaic virus (strain Strasbourg) (CamV).
OC Viruses; Retroid viruses; Caulimoviridae; Caulimovirus.
OX NCBI_TaxID=10648;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81001865; PubMed=7407912;
RA Franck A., Guille H., Jonard G., Richards K., Hirth L.;
RT "Nucleotide sequence of cauliflower mosaic virus DNA."
RL Cell 21:285-294(1980).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate
CC + [DNA](N).
CC -1- SIMILARITY: HIGH, WITH OTHER CAULIMOVIRUS ORF V.
CC -1- SIMILARITY: WITH RETROVIRAL POL/GAG POLYPEPTIDES.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V00141; CAA23460.1; -.
DR PIR: A04159; QOCV5.
DR MEROPS: A03.001; -.
DR InterPro: IPR000588; Peptidase_A3.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; RVT; 1.
DR Pfam: PF02160; Peptidase_A3; 1.
DR PRINTS: PR00731; CAULIMOPASE.
KW Hydrolyase; Aspartyl protease; RNA-directed DNA polymerase;
KW Transferase; Endonuclease; Polypeptide.
FT ACT_SITE 45 45 BY SIMILARITY.
FT SIMILAR 40 130 TO RETROVIRUS GAG/POL PROTEASE DOMAIN.
FT SIMILAR 260 620 TO RETROVIRUS GAG/POL DNA POLYMERASE
FT DOMAIN.
SQ SEQUENCE 679 AA; 78629 MM; 9EE527BCD460B766 CRC64;

Query Match 2.4%; Score 7; DB 1; Length 679;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 223 LKVIKPS 229
    |||||
Db 272 LKVIKPS 278

RESULT 42
POL_CAMVN STANDARD; PRT; 680 AA.
AC 000962;
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enzymatic polypeptide [Contains: Aspartic protease (EC 3.4.23.-);
DE Endonuclease; Reverse transcriptase (EC 2.7.7.49)].
GN V.
OS Cauliflower mosaic virus (strain NY8153) (CamV).
OC Viruses; Retroid viruses; Caulimoviridae; Caulimovirus.
OX NCBI_TaxID=31557;
RN [1]
RP SEQUENCE FROM N.A.
RA Chenault K.D., Steffens D.L., Melcher U.K.;
RT "Nucleotide sequence of cauliflower mosaic virus isolate NY8153."
RL Plant Physiol. 100:542-545(1992).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate
CC + [DNA](N).
CC -1- SIMILARITY: HIGH, WITH OTHER CAULIMOVIRUS ORF V.
CC -1- SIMILARITY: WITH RETROVIRAL POL/GAG POLYPEPTIDES.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: M9541; AAA46358.1; -
 CC
 CC DR MEROPS; A03.001; -
 CC DR InterPro: IPR000586; Peptidase_A3.
 CC DR InterPro: IPR000477; RTase.
 CC DR Pfam: PF00078; rvc.1.
 CC DR Pfam: PF02160; Peptidase_A3.1.
 CC DR PRINTS: PRO0731; CAULIMOPTASE.
 CC DR Hydroxylase: Aspartyl protease; RNA-directed DNA polymerase;
 CC KM Transferase; Endonuclease; Polypeptide.
 CC FT ACT_SITE 46 46 BY SIMILARITY.
 CC FT SIMILAR 41 131 TO RETROVIRUS GAG/POL PROTEASE DOMAIN.
 CC FT SIMILAR 261 621 TO RETROVIRUS GAG/POL DNA POLYMERASE
 CC FT SIMILAR 261 621 DOMAIN.
 CC SQ SEQUENCE 680 AA; 78665 MW; FCE02E09647C2221 CRC64;

Query Match 2.4%; Score 7; DB 1; Length 680;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 223 LKVIKPS 229
 |||||
 DB 273 LKVIKPS 279

RESULT 43
 VARI_METUA STANDARD; PRT; 695 AA.
 ID VARI_METUA
 AC 057675;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE V-type ATP synthase subunit I (EC 3.6.3.14) (V-type ATPase subunit I).
 GN ATP1 OR MJ0222.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE-96337999; PubMed-8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.T.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uitterlinden T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Bordovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.
 RL Science 273:1058-1073(1996).
 CC
 CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
 CC GRADIENT ACROSS THE MEMBRANE.
 CC
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
 CC H(+) (out).
 CC
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U67478; AAB98208.1; -
 DR TIGR: MJO222; -
 DR InterPro: IPR002490; V_ATPase_sub116.
 DR Pfam: PF01496; V_ATPase_sub_a.1.
 DR Hydroxylase; hydrogen ion transport; Transmembrane; Complete proteome.
 KM TRANSMEM 392 412 POTENTIAL.
 FT TRANSMEM 425 445 POTENTIAL.
 FT TRANSMEM 497 517 POTENTIAL.
 FT TRANSMEM 534 554 POTENTIAL.
 FT TRANSMEM 556 576 POTENTIAL.
 FT TRANSMEM 605 625 POTENTIAL.
 FT TRANSMEM 627 647 POTENTIAL.
 FT TRANSMEM 627 647 POTENTIAL.
 SQ SEQUENCE 695 AA; 76953 MW; F03B5CEBE29D53D CRC64;

Query Match 2.4%; Score 7; DB 1; Length 695;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 218 SLIEKTK 224
 |||||
 DB 266 SLIEKTK 272

RESULT 44
 ADHE_CIOAB STANDARD; PRT; 863 AA.
 ID ADHE_CIOAB
 AC P33744; Q45809; Q45808;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Aldehyde-alcohol dehydrogenase [includes: Alcohol dehydrogenase
 DE (EC 1.1.1.1) (ADH); Acetaldehyde dehydrogenase [acetylating]
 DE (EC 1.2.1.10) (ACDH).
 GN ADHE OR AAD OR CAP0162.
 OS Clostridium acetobutylicum.
 OC Clostridia; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE-94042861; PubMed-8226639;
 RA Fischer R.J., Helms J., Duerre P.;
 RT Cloning, sequencing, and molecular analysis of the sol operon of
 RT Clostridium acetobutylicum, a chromosomal locus involved in
 RT solventogenesis.
 RL J. Bacteriol. 175:6959-6969(1993).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE-9411967; PubMed-8300540;
 RA Nair R.V., Bennett G.N., Papoutsakis E.T.;
 RT Molecular characterization of an aldehyde/alcohol dehydrogenase gene
 RT from Clostridium acetobutylicum ATCC 824.
 RL J. Bacteriol. 176:871-885(1994).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE-21359325; PubMed-11466286;
 RA Neelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitt J., Wolf Y.I.,
 RA Tatunov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum.
 RL J. Bacteriol. 183:4823-4838(2001).
 CC
 CC -1- FUNCTION: THIS ENZYME HAS PROBABLY TWO ACTIVITIES: ADH, AND ACDH.
 CC
 CC -1- CATALYTIC ACTIVITY: An alcohol + NAD(+) = an aldehyde or ketone +
 CC NADH.
 CC
 CC -1- INDUCTION: INDUCED DURING SWITCH TO SOLVENT PRODUCTION.

```

CC -1- SIMILARITY: TO THE C. KLUYVERI NADP-DEPENDENT SUCCINATE
CC SEMIALDEHYDE DEHYDROGENASE (SUCC) (IN THE N-TERMINAL SECTION).
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION: BELONGS TO THE IRON-
CC CONTAINING ALCOHOL DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X72831; CAA51344.1; -
DR EMBL; L14817; AAD04638.1; -
DR EMBL; AE001438; AAK76907.1; -
DR PIR; S33434; S33434.
DR PIR; A49346; A49346.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR001670; Fe-ADH.
DR Pfam; PF00171; aldehyd; 1.
DR PROSITE; PS00913; ADH_IRON_1; 1.
DR PROSITE; PS00060; ADH_IRON_2; FALSE_NEG.
DR Oxidoreductase; Multifunctional enzyme; NAD; Plasmid;
KM Complete proteome.
FT ACT_SITE 244 244 BY SIMILARITY.
FT NP_BIND 420 425 NAD (POTENTIAL).
SQ SEQUENCE 862 AA; 95321 MW; BE09E32B8DD08B0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 862;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 EKLKVIK 227
   |||||
Db 11 EKLKVIK 17

RESULT 45
YML7_YEAST STANDARD; PRT; 1134 AA.
ID YML7_YEAST
AC Q03735;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 126.1 kDa protein in ND11-ATRI Intergenic region.
GN YML117W OR YM8339.02.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Skellton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -1- SIMILARITY: TO YEAST YPL184C AND S.POMBE NRD1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z49210; CAA89101.1; -
DR SGD; S0004585; YML117W
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR003955; RRM_2.
DR SMART; SM00360; RRM; 1.

```

```

DR SMART; SM00362; RRM_2; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KM Hypothetical protein; RNA-binding.
FT DOMAIN 653 726
SQ SEQUENCE 1134 AA; 126138 MW; 47B0FF7C5B98A053 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 1134;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 IDLHSIV 40
   |||||
Db 225 IDLHSIV 231

Search completed: May 21, 2003, 11:12:15
Job time : 16 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 10:58:41 ; Search time 33 Seconds
(without alignments)
1791.986 Million cell updates/sec

Title: US-09-869-677A-2

Perfect score: 287
Sequence: 1 SGTGAKTKAKSKLKVAVTNS.....PQDSYYAMKKMLDKISGL 287

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL_21:*

1: sp.archaea:*\n2: sp.bacteria:*\n3: sp.fungi:*\n4: sp.human:*\n5: sp.invertebrate:*\n6: sp.mammal:*\n7: sp.mhc:*\n8: sp.organelle:*\n9: sp.phage:*\n10: sp.plant:*\n11: sp.rodent:*\n12: sp.virus:*\n13: sp.vertebrate:*\n14: sp.unclassified:*\n15: sp.virus:*\n16: sp.bacteriap:*\n17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	27	9.4	309	2	Q8V082	Q8V082 streptococ
2	26	9.1	80	2	Q9RIR5	Q9RIR5 streptococ
3	26	9.1	81	2	Q9RIL0	Q9RIL0 streptococ
4	11	3.8	308	2	Q47723	Q47723 enterococ
5	11	3.8	309	2	Q9ZFP2	Q9ZFP2 enterococ
6	9	3.1	292	2	Q9ZT13	Q9ZT13 enterococ
7	9	3.1	313	2	Q8VNI4	Q8VNI4 lactobacill
8	8	2.8	255	16	Q9HZS1	Q9HZS1 pseudomonas
9	8	2.8	269	5	Q9VRS0	Q9VRS0 drosophila
10	8	2.8	309	2	P72415	P72415 staphylococ
11	8	2.8	309	16	Q99VY4	Q99VY4 staphylococ
12	8	2.8	312	2	Q8VQS9	Q8VQS9 staphylococ
13	8	2.8	353	16	Q8R8V5	Q8R8V5 thermoaer
14	8	2.8	428	16	Q34862	Q34862 bacillus su
15	8	2.8	618	17	Q8U211	Q8U211 pyrococcus
16	8	2.8	3072	12	Q92645	Q92645 clover yell

17	7	2.4	54	3	P79070	P79070 debaryomyce
18	7	2.4	80	2	Q9FR42	Q9FR42 mycobacteri
19	7	2.4	90	16	Q8RE73	Q8RE73 agrobacteri
20	7	2.4	109	10	Q93W30	Q93W30 arabidopsis
21	7	2.4	112	16	Q92HD5	Q92HD5 rickettsia
22	7	2.4	114	12	Q9UG75	Q9UG75 tt virus. o
23	7	2.4	114	12	Q9UG72	Q9UG72 tt virus. o
24	7	2.4	114	12	Q9UG69	Q9UG69 tt virus. o
25	7	2.4	114	12	Q9UG66	Q9UG66 tt virus. o
26	7	2.4	124	2	Q8RT58	Q8RT58 synchococ
27	7	2.4	125	17	Q9HJ05	Q9HJ05 thermoplasm
28	7	2.4	135	16	Q8X523	Q8X523 escherichia
29	7	2.4	142	5	Q61985	Q61985 caenorhabd
30	7	2.4	148	4	Q96GS5	Q96GS5 homo sapien
31	7	2.4	150	11	Q9D3D7	Q9D3D7 mus musculu
32	7	2.4	155	2	Q9ZG16	Q9ZG16 bordetella
33	7	2.4	155	2	Q9S3M4	Q9S3M4 bordetella
34	7	2.4	155	5	P91468	P91468 caenorhabd
35	7	2.4	155	16	Q97KFO	Q97KFO clostridium
36	7	2.4	161	11	Q8YEA3	Q8YEA3 mus musculu
37	7	2.4	162	16	Q8ZPG2	Q8ZPG2 salmonella
38	7	2.4	162	16	Q8ZJ78	Q8ZJ78 salmonella
39	7	2.4	170	2	Q9ZJ37	Q9ZJ37 helicobacte
40	7	2.4	172	16	Q8UC67	Q8UC67 agrobacteri
41	7	2.4	184	5	Q16158	Q16158 drosophila
42	7	2.4	192	16	Q9Z8B7	Q9Z8B7 chlamydia p
43	7	2.4	192	16	Q8TFN3	Q8TFN3 bruceella me
44	7	2.4	195	11	Q9D920	Q9D920 mus musculu
45	7	2.4	196	4	Q969J3	Q969J3 homo sapien

ALIGNMENTS

RESULT 1

Q8V082 Q8V082 PRELIMINARY; PRT; 309 AA.
AC Q8V082;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pneumococcal surface adhesin A.
GN PSA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-KN1H2033;
RA Bae S.M., Lee K.J.;
RT "Molecular characterization of a gene coding for the pneumococcal surface adhesin A (PSA) of Streptococcus pneumoniae isolated from the patients in Korea.";
RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF459638; AAL65356.1; -;
DR EMBL: AF459638; AAL65356.1; -;
DR InterPro: IPR001987; Lipoprotein_4.
DR Pfam: PF01297; SBP_bac_9.1.
DR PRINTS: PR00690; ADHESINFAMILY.
SQ SEQUENCE 309 AA; 34522 MW; B125E7FE3EC6F67C CRC64;

Query Match 9.4%; Score 27; DB 2; Length 309;
Best Local Similarity 100.0%; Pred. No. 4.4e-19;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 IADKDKLHSLVPIGDPHEPEPLPD 55
|||||
Db 49 IADKDKLHSLVPIGDPHEPEPLPD 75

RESULT 2
Q9RIR5 Q9RIR5 PRELIMINARY; PRT; 80 AA.

```

AC 09RT5:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative metal binding/adhesin protein (Fragment).
GN MBA.
OS Streptococcus equi.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCTC 9682;
RX MEDLINE=20448974; PubMed=10992520;
RA Harrington D.J., Greated J.S., Chanter N., Sutcliffe I.C.;
RT "Identification of Lipoprotein Homologues of Pneumococcal PsaA in the
RT Equine Pathogens Streptococcus equi and Streptococcus zooepidemicus.";
RL Infect. Immun. 68:6048-6051(2000).
DR EMBL: AJ249889; CAB56709.2;
DR InterPro: IPR001987; Lipoprotein_4.
DR Pfam: PF01297; SBP_bac_9; 1.
FT NON_TER
SQ SEQUENCE 80 AA; 9081 MW; 94B40C7278E25760 CRC64;

Query Match
Best Local Similarity 9.1%; Score 26; DB 2; Length 80;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 177 NKLIIVTSGCFKFSKAYGVPASAI 202
DB 55 NKLIIVTSGCFKFSKAYGVPASAI 80

RESULT 3
O9RILO
ID O9RILO PRELIMINARY; PRT; 81 AA.
AC O9RILO.
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative metal binding/adhesin protein (Fragment).
GN MBA.
OS Streptococcus equi subsp. zooepidemicus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=40041;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCTC 7023;
RA Harrington D.J., Chanter N., Greated J.S., Sutcliffe I.C.;
RT "Identification of homologues of the pneumococcal PsaA protein in the
RT equine pathogens Streptococcus equi and Streptococcus zooepidemicus.";
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ249890; CAB56710.1;
DR InterPro: IPR001987; Lipoprotein_4.
DR Pfam: PF01297; SBP_bac_9; 1.
FT NON_TER
FT NON_TER
SQ SEQUENCE 81 AA; 9195 MW; EBA1264D6DC81760 CRC64;

Query Match
Best Local Similarity 9.1%; Score 26; DB 2; Length 81;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 177 NKLIIVTSGCFKFSKAYGVPASAI 202
DB 56 NKLIIVTSGCFKFSKAYGVPASAI 81

RESULT 4
O47723
ID O47723 PRELIMINARY; PRT; 308 AA.

```

```

AC Q47723:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Endocarditis specific antigen.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95122212; PubMed=7822045;
RA Lowe A.W., Lambert P.A., Smith A.W.;
RT "Cloning of an Enterococcus faecalis endocarditis antigen: homology
RT with adhesins from some oral streptococci.";
RL Infect. Immun. 63:703-706(1995).
DR EMBL: U03756; AA70056.1;
DR InterPro: IPR001987; Lipoprotein_4.
DR Pfam: PF01297; SBP_bac_9; 1.
DR PRINTS: PRO0690; ADHESNFAMILY.
SQ SEQUENCE 308 AA; 34768 MW; 2EC868BF34EF276 CRC64;

Query Match
Best Local Similarity 3.8%; Score 11; DB 2; Length 308;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 DPHEYEPLPED 55
DB 63 DPHEYEPLPED 73

RESULT 5
O9ZFP2
ID O9ZFP2 PRELIMINARY; PRT; 309 AA.
AC O9ZFP2.
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Efma (Fragment).
GN EFMA.
OS Enterococcus faecium (Streptococcus faecium).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1352;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UOI;
RA Flatman J.C., Smith A.W.;
RT "Identification of a putative ABC Operon in Enterococcus faecium.";
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF097414; AAC68699.1;
DR InterPro: IPR001987; Lipoprotein_4.
DR Pfam: PF01297; SBP_bac_9; 1.
DR PRINTS: PRO0690; ADHESNFAMILY.
FT NON_TER
FT NON_TER
SQ SEQUENCE 309 AA; 34538 MW; 9269BEC376BD64BD CRC64;

Query Match
Best Local Similarity 3.8%; Score 11; DB 2; Length 309;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 DPHEYEPLPED 55
DB 65 DPHEYEPLPED 75

RESULT 6
O9Z131
ID O9Z131 PRELIMINARY; PRT; 292 AA.
AC O9Z131.
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

```

DE Efaim.
 GN EFAM.
 OS Enterococcus faecium (Streptococcus faecium).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 CC Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1352;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TX1330;
 RX MEDLINE=98423927; PubMed=9753005;
 RA Singh K.V., Coque T.M., Weinstock G.M., Murray B.E.;
 RT "In vivo testing of an Enterococcus faecalis eta mutant and use of
 eta homologs for species identification."
 RL FEMS Immunol. Med. Microbiol. 21:323-331(1998).
 DR EMBL; AF042288; AAC64251.1; -;
 DR InterPro: IPR001987; LipoProtein_4.
 DR Pfam: PF01297; SBP_dac_9; 1.
 DR PRINTS: PR00690; ADHESNFAMILY.
 SQ SEQUENCE 292 AA; 33052 MW; 1BB51038102099BC CRC64;

Query Match 3.1%; Score 9; DB 2; Length 292;
 Best Local Similarity 100.0%; Pred. No. 1.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 200 AYIWEINTE 208
 Db 208 AYIWEINTE 216

RESULT 7
 OSVN14 PRELIMINARY; PRT; 313 AA.
 AC Q8VN14;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Substrate binding protein.
 GN MWS.
 OS Lactobacillus casei.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 CC Lactobacillaceae; Lactobacillus.
 OX NCBI_TaxID=1582;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DSM 20011;
 RA Nietop-groot M.N., Pentcheva E., Verdoes J.C.;
 RT "Manganese transport in Lactobacillus casei."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ276708; CAC81445.1; -;
 DR InterPro: IPR001987; LipoProtein_4.
 DR Pfam: PF01297; SBP_dac_9; 1.
 DR PRINTS: PR00690; ADHESNFAMILY.
 SQ SEQUENCE 313 AA; 35097 MW; A9A45B546B2E9A85 CRC64;

Query Match 3.1%; Score 9; DB 2; Length 313;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 200 AYIWEINTE 208
 Db 224 AYIWEINTE 232

RESULT 8
 O9HZS1 PRELIMINARY; PRT; 255 AA.
 AC O9HZS1;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Histidine transport protein HISP.
 GN HISP OR PA2926.
 OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Ervin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Huynh W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Salter M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen."
 RL Nature 406:959-964(2000).
 CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
 (ABC TRANSPORTERS).
 DR EMBL; AE004718; AAG06314.1; -;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00005; ABC_tran; 1.
 DR ProDom: PD000006; ABC_transport; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 DR ATP-binding; Transport; Complete proteome.
 SQ SEQUENCE 255 AA; 28518 MW; 3B8F587FCD91892B CRC64;

Query Match 2.8%; Score 8; DB 16; Length 255;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 208 EEEGTPDQ 215
 Db 229 EEEGTPDQ 236

RESULT 9
 OSVN30 PRELIMINARY; PRT; 269 AA.
 AC Q9VR30;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE CG15628 protein.
 GN CG15628.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
 CC Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 DE Ephydroidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jajali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA SVIRSKAS R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003575; AAF50978.2; -
 DR FlyBase: FBgn0031632; CG15628.
 DR InterPro: IPR000182; GCSAcetyltransf.
 DR Pfam: PF00583; Acetyltransf.1.
 SQ SEQUENCE 269 AA; 30719 MW; B5FEE31B0189B252 CRC64;

Query Match 2.8%; Score 8; DB 5; Length 269;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 50 EPLPEDAE 57
 |||||||
 Db 58 EPLPEDAE 65

RESULT 10
 P72415 PRELIMINARY; PRT; 309 AA.
 AC P72415;
 DT 01-FEB-1997 (TREMblrel. 02, Created)
 DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Lipoprotein.
 OS *Staphylococcus epidermidis*.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 CC *Staphylococcus*.
 OX NCBI_TaxID=1282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=901;
 RX MEDLINE=98339879; PubMed=9673260;
 RA Cockayne A., Hill P.J., Powell N.B.L., Bishop K., Sims C.,
 RA Williams P.;
 RT "Molecular cloning of a 32-kilodalton lipoprotein component of a novel
 RT iron-regulated *Staphylococcus epidermidis* ABC transporter."
 RL Infect. Immun. 66:3767-3774(1998).
 DR EMBL: X99127; CA67571.1; -
 DR InterPro: IPR001987; Lipoprotein_4.
 DR Pfam: PF01297; SBP_bac_9; 1.
 DR PRINTS: PR00690; ADHESNFAMILY.
 SQ SEQUENCE 309 AA; 34832 MW; D91E934A901FA46 CRC64;

Query Match 2.8%; Score 8; DB 2; Length 309;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 GQDPHEYE 50
 |||||||
 Db 63 GQDPHEYE 70

RESULT 11
 O99VY4

ID O99VY4 PRELIMINARY; PRT; 309 AA.
 AC O99VY4;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE SA0587 protein (Lipoprotein).
 GN SAV0631 OR SA0587.
 OS *Staphylococcus aureus* (strain Mu50 / ATCC 706699), and
 OS *Staphylococcus aureus* (strain N315).
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 CC *Staphylococcus*.
 OX NCBI_TaxID=158879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-S.aureus (strain Mu50), and S.aureus (strain N315);
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mitutani O.I.Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hiraoka H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiratake K.;
 RT "Whole genome sequencing of methicillin-resistant *Staphylococcus*
 RT aureus."
 RL Lancet 357:1225-1240(2001).
 DR EMBL: AP003360; BAB56793.1; -
 DR EMBL: AP003131; BAB41819.1; -
 DR InterPro: IPR001987; Lipoprotein_4.
 DR Pfam: PF01297; SBP_bac_9; 1.
 DR PRINTS: PR00690; ADHESNFAMILY.
 DR Complete Proteome.
 SQ SEQUENCE 309 AA; 34741 MW; 3BE524E31D76AF04 CRC64;

Query Match 2.8%; Score 8; DB 16; Length 309;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 GQDPHEYE 50
 |||||||
 Db 63 GQDPHEYE 70

RESULT 12
 O8VOS9 PRELIMINARY; PRT; 312 AA.
 ID O8VOS9
 AC O8VOS9;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE MntC.
 GN MNTC.
 OS *Staphylococcus aureus*.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 CC *Staphylococcus*.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8325-4;
 RA Horsburgh M.T., Wharton S.J., Cox A.G., Ingham E., Peacock S.,
 RA Foster S.J.;
 RT "NtrC modulates expression of the ferr regulon and SOD-independent
 RT superoxide resistance in *Staphylococcus aureus* through control of
 RT manganese uptake."
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF452622; AAL50778.1; -
 DR InterPro: IPR001987; Lipoprotein_4.
 DR Pfam: PF01297; SBP_bac_9; 1.
 DR PRINTS: PR00690; ADHESNFAMILY.
 SQ SEQUENCE 312 AA; 35071 MW; 743BA4EADC36A7CE CRC64;

Query Match 2.8%; Score 8; DB 2; Length 312;
 Best Local Similarity 100.0%; Pred. No. 15;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 GODPHEYE 50
|||||||

Db 63 GODPHEYE 70

RESULT 13

O8R8V5 PRELIMINARY; PRT; 353 AA.

AC O8R8V5
DT 01-JUN-2002 (TRENBLREL. 21, Created)
DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
DE 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
GN BIOB2 OR TTE1885.
OS Thermomicrobacter tengcongensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Thermomicrobacteriales; Thermomicrobacteriaceae; Thermomicrobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MBAT / JCM11007;
RA MEDLINE=2192816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Huang H.;
RT "A complete sequence of T. tengcongensis genome."
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013140; AAM25069.1; -.
KW Complete proteome.
SQ SEQUENCE 353 AA; 40112 MW; CA9D7DD0A8C03811 CRC64;

Query Match 2.8%; Score 8; DB 16; Length 353;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 216 ISSLEKTL 223
|||||||

Db 11 ISSLEKTL 18

RESULT 14

O34862 PRELIMINARY; PRT; 428 AA.

ID O34862
AC O34862;
DT 01-JAN-1998 (TRENBLREL. 05, Created)
DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE YTCA.
GN YTCA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA Lapidus A., Galleron N., Sorokin A., Ehrlich D.,
RA Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertolo M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignelli S.C., Bron S.,
RA Bouilliet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Gutseppl G., Guy B.D., Haga K., Halech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianchand M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
RA Rieger M., Rivolta C., Kocha E., Koche B., Rose M., Sedate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpsira P., Tognoni A.,
RA Tostato V., Uchiyama S., Vandenbol M., Vannier F., Vassaretti A.,
RA Viati A., Wandt R., Wedler E., Wedler K., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis".
RL Nature 390:249-256(1997).

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.,
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF008220; AAC00367.1; -.
DR EMBL; Z99119; CAB15064.1; -.
DR InterPro; IPR00205; NAD_binding.
DR InterPro; IPR001732; UDPG_MGDP_dh.
DR Pfam; PF00984; UDPG_MGDP_dh; 1.
DR Pfam; PF03720; UDPG_MGDP_dh; 1.
DR Pfam; PF03721; UDPG_MGDP_dh; 1.
KW Complete proteome.
SQ SEQUENCE 428 AA; 46832 MW; 70A68AAE11716384 CRC64;

Query Match 2.8%; Score 8; DB 16; Length 428;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 NIAKOLIA 140
|||||||

Db 126 NIAKOLIA 133

RESULT 15

O80211 PRELIMINARY; PRT; 618 AA.

ID O80211
AC O80211;
DT 01-JUN-2002 (TRENBLREL. 21, Created)
DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Hypothetical protein PF1041.
GN PF1041.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010215; AAL81165.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 618 AA; 70210 MW; BE8562691A6D566 CRC64;

Query Match 2.8%; Score 8; DB 17; Length 618;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 101 GIDVYLE 108

Db 137 GIDVTYPE 144

RESULT 16

092645 PRELIMINARY; PRT; 3072 AA.

AC 092645; (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Polypeptide.
 OS Clover yellow vein virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 CC Potyvirus.
 CX NCBI_TaxID=12198;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NO.30;
 RA Uyeda I., Takahashi Y.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NO.30;
 RA MEDLINE-97456254; PubMed-9311568;
 RA Takahashi Y., Takahashi T., Uyeda I.;
 RT "A cDNA clone to clover yellow vein potyvirus genome is highly
 infectious."
 RT Virus Genes 14:235-243(1997).
 RL EMBL; AB011819; BAA25147.1; -.
 DR MEROPS; C04.008; -.
 DR MEROPS; C06.001; -.
 DR InterPro; IPR001410; Helicase_C.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR001730; Peptidase_C4.
 DR InterPro; IPR001456; Peptidase_C6.
 DR InterPro; IPR001592; Poly_coat.
 DR InterPro; IPR002540; Poly_p1.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR001254; Ser.protease_Try.
 DR Pfam; PF00271; helicase_C; 1.
 DR Pfam; PF00863; peptidase_C4; 1.
 DR Pfam; PF00831; peptidase_C6; 1.
 DR Pfam; PF00767; Poly_coat; 1.
 DR Pfam; PF01577; Poly_p1; 1.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR PRINTS; PR00966; N1APOTYPASE.
 DR SMART; SMO0487; DEXDC; 1.
 DR SMART; SMO0480; HELICG; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR ATP-binding; Helicase; Hydrolase; Serine protease.
 KW SEQUENCE 3072 AA; 348727 MW; F3BEA84BD96D385A CRC64;

Query Match 2.8%; Score 8; DB 12; Length 3072;

Best Local Similarity 100.0%; Pred. No. 1.2e+02; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 221 EKLKVIKP 228

Db 149 EKLKVIKP 156

RESULT 17

P79070 PRELIMINARY; PRT; 54 AA.

AC P79070;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Reverse transcriptase (Fragment).
 OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.

OX NCBI_TaxID=4959;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-IGC 2968;
 RA Roberts I.N.;
 RT "Reverse transcriptase gene sequences in Debaryomyces hansenii and
 Dekkera anomala."
 RT Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y11869; CAA72617.1; -.
 FT NON_TER 1 1
 FT NON_TER 54 54
 SQ SEQUENCE 54 AA; 6311 MW; 1995703BE4959DA3 CRC64;

Query Match 2.4%; Score 7; DB 3; Length 54;

Best Local Similarity 100.0%; Pred. No. 32;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 176 ENKRLIV 182

Db 40 ENKRLIV 46

RESULT 18

Q9F842 PRELIMINARY; PRT; 80 AA.

AC Q9F842;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical methyltransferase (Fragment).
 OS Mycobacterium smegmatis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CX NCBI_TaxID=1772;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-20374587; PubMed-10913817;
 RA Mukhopadhyay B., Purwantinili E.;
 RT "Pyruvate carboxylase from Mycobacterium smegmatis: stabilization,
 RT rapid purification, molecular and biochemical characterization and
 regulation of the cellular level."
 RT Blochim. Biophys. Acta 1475:191-206(2000).
 RL EMBL; AF262949; AAG30412.1; -.
 DR InterPro; IPR004398; Cons_hypoth95.
 DR Pfam; PF03602; Cons_hypoth95; 1.
 DR Pfam; PF03602; Cons_hypoth95; 1.
 KW Methyltransferase; Transferase.
 FT NON_TER 80 80
 SQ SEQUENCE 80 AA; 8307 MW; 0B89C7F49D6FEAD7 CRC64;

Query Match 2.4%; Score 7; DB 2; Length 80;

Best Local Similarity 100.0%; Pred. No. 46;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 229 SALFVES 235

Db 68 SALFVES 74

RESULT 19

O8UEF3 PRELIMINARY; PRT; 90 AA.

AC O8UEF3;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein Atu1805.
 GN Atu1805 OR AGR_C.3318.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE-21608550; PubMed-11743193;
 RA Okuda D.M., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Wood V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Moo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. St.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.M.,
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58.";
 RL Science 294:2317-2323(2001).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21608551; PubMed-11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Querillo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
 RA Houmlet K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.,
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL: AE009135; AAL42803.1; -;
 DR EMBL: AE008101; AAK87575.1; -;
 KW Hypothetical protein, complete proteome.
 SQ SEQUENCE 90 AA; 9824 MW; C2FC7AD8355F027 CRC64;

Query Match 2.4%; Score 7; DB 16; Length 90;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 103 DVIYLEG 109
 |||||
 Db 53 DVIYLEG 59

RESULT 20
 OY3W30 PRELIMINARY; PRT; 109 AA.
 AC Q93W30;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Hypothetical 12.0 kDa protein (Expressed protein).
 GN AT2G35605.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamuya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Koeseima E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
 RA Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene At2g35605 (not previously annotated).";
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K., Yu G.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C.C., Toriumi M., Yu G.,
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamuya A., Karlin-Neumann G., Kawai J., Kim C., Koeseima E.,
 RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,

RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene At2g35605 (not previously annotated).";
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
 RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
 RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Town C.D., Kaul S.;
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., Yanaken S.E.,
 RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
 RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
 RA Fraser C.M., Venter J.C.;
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY059124; AAL15230.1; -;
 DR EMBL: AF370234; AAK44049.1; -;
 DR EMBL: AC005314; AAM15040.1; -;
 DR EMBL: AC006068; AAM15113.1; -;
 DR InterPro: IPR003121; SWIB.
 DR Pfam: PF02201; SWIB; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 109 AA; 11951 MW; 5F63922DC1467A8 CRC64;

Query Match 2.4%; Score 7; DB 10; Length 109;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SSTGAKT 7
 |||||
 Db 17 SSTGAKT 23

RESULT 21
 OY2HD5 PRELIMINARY; PRT; 112 AA.
 AC Q92HD5;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Similarity to 3-hydroxyacyl-CoA dehydrogenase (Fadh).
 GN RC0836.
 OS Rickettsia conorii.
 CC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 CC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=781;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MALISH 7;
 RA MEDLINE-21442074; PubMed-11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
 RL Science 293:2093-2098(2001).
 DR EMBL: AE008639; AAL03374.1; -;
 KW Complete proteome.
 SQ SEQUENCE 112 AA; 13051 MW; 0C2E3322BA605CB6 CRC64;

Query Match 2.4%; Score 7; DB 16; Length 112;

Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 IERLKV1 226
|||||
Db 104 IERLKV1 110

RESULT 22

Q9UG75 PRELIMINARY; PRT; 114 AA.
AC Q9UG75;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ORF3.
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TIVON-KC009;
RX MEDLINE-20428649; PubMed-10971131;
RA Takahashi K., Hijikata M., Samokhvalov E.I., Mishiro S.;
RT "Full or near full length nucleotide sequences of TT virus variants
(types SANBAN and YONBAN) and the TT virus-like mini virus.";
RL Intervirology 43:119-123(2000).
DR EMBL: AB038621; BAA93584.1;
SQ SEQUENCE 114 AA; 12597 MW; FE45E1739405F8BA CRC64;

Query Match 2.4%; Score 7; DB 12; Length 114;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AOKTKNK 93
|||||
Db 45 AOKTKNK 51

RESULT 23

Q9UG72 PRELIMINARY; PRT; 114 AA.
AC Q9UG72;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ORF3.
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TIVON-KC011;
RX MEDLINE-20428649; PubMed-10971131;
RA Takahashi K., Hijikata M., Samokhvalov E.I., Mishiro S.;
RT "Full or near full length nucleotide sequences of TT virus variants
(types SANBAN and YONBAN) and the TT virus-like mini virus.";
RL Intervirology 43:119-123(2000).
DR EMBL: AB038622; BAA93587.1;
SQ SEQUENCE 114 AA; 12574 MW; 103E0820E834AE92 CRC64;

Query Match 2.4%; Score 7; DB 12; Length 114;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AOKTKNK 93
|||||
Db 45 AOKTKNK 51

RESULT 24
Q9UG69 PRELIMINARY; PRT; 114 AA.

AC Q9UG69;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ORF3.
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TIVON-KC186;
RX MEDLINE-20428649; PubMed-10971131;
RA Takahashi K., Hijikata M., Samokhvalov E.I., Mishiro S.;
RT "Full or near full length nucleotide sequences of TT virus variants
(types SANBAN and YONBAN) and the TT virus-like mini virus.";
RL Intervirology 43:119-123(2000).
DR EMBL: AB038623; BAA93590.1;
SQ SEQUENCE 114 AA; 12645 MW; 7D01B8083F44B27 CRC64;

Query Match 2.4%; Score 7; DB 12; Length 114;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AOKTKNK 93
|||||
Db 45 AOKTKNK 51

RESULT 25

Q9UG66 PRELIMINARY; PRT; 114 AA.
AC Q9UG66;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ORF3.
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TIVON-KC197;
RX MEDLINE-20428649; PubMed-10971131;
RA Takahashi K., Hijikata M., Samokhvalov E.I., Mishiro S.;
RT "Full or near full length nucleotide sequences of TT virus variants
(types SANBAN and YONBAN) and the TT virus-like mini virus.";
RL Intervirology 43:119-123(2000).
DR EMBL: AB038624; BAA93593.1;
SQ SEQUENCE 114 AA; 12700 MW; D2162CE73FA52BE CRC64;

Query Match 2.4%; Score 7; DB 12; Length 114;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 AOKTKNK 93
|||||
Db 45 AOKTKNK 51

RESULT 26

Q8RT58 PRELIMINARY; PRT; 124 AA.
AC Q8RT58;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cyanoglobin.
GN GLEN.
OS Synechococcus sp. (strain PCC 7002) (Agmenellum quadruplicatum).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32049;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-PCC 7002;
 RA Maguadadt J., Bryant D.A., Zhao J., Li T.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF475938; AAL79195.1; -;
 SO SEQUENCE 124 AA; 13857 MW; 8B7CB4D36132E287 CRC64;

Query Match 2.4%; Score 7; DB 2; Length 124;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 171 FDAIEN 177
 DB 84 FDAIEN 90

RESULT 27

ID 09HJ05 PRELIMINARY; PRT; 125 AA.
 AC 09HJ05;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DR InterPro: IPR004701; EITa-man.
 DE 01-JUN-2002 (TREMblrel. 16, Last sequence update)
 DE Hypothetical protein Ta0868.
 GN TA0868.
 OS Thermoplasma acidophilum.
 OC Archaea: Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 CC Thermoplasmataceae; Thermoplasmata.
 OX NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DSM 1728;
 RX MEDLINE=20479972; PubMed-11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RT Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermophilic scavenger Thermoplasma
 acidophilum.";
 RL Nature 407:508-513(2000).
 DR EMBL: AL445065; CAC11997.1; -;
 DR InterPro: IPR005149; Pdr.
 DR Pfam: PF03551; Pdr; 1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 125 AA; 14043 MW; 82EDEDC157858E8F CRC64;

Query Match 2.4%; Score 7; DB 17; Length 125;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 111 SEKRED 117
 DB 49 SEKRED-55

RESULT 28

ID 08X523 PRELIMINARY; PRT; 135 AA.
 AC 08X523;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DR 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Putative sorbose PTS component (sorbose-permease PTS system IIA
 component).
 GN Z5617 OR ECSS002.
 OS Escherichia coli O157:H7.
 CC Bacteria: Proteobacteria; gamma subdivision: Enterobacteriaceae;
 CC Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed-11205551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,

RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 RN [2]

Query Match 2.4%; Score 7; DB 16; Length 135;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 VAKLEKL 163
 DB 44 VAKLEKL 50

RESULT 29

ID 061985 PRELIMINARY; PRT; 142 AA.
 AC 061985;
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DR 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 GN K02H11.6.
 DE Hypothetical 16.2 kDa protein.
 OS Caenorhabditis elegans.
 CC Eukaryota: Metazoa: Nematoda: Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed-9851916;
 RA Waterston R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]

Query Match 2.4%; Score 7; DB 5; Length 142;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 VAKLEKL 163
 DB 44 VAKLEKL 50

RESULT 29

ID 061985 PRELIMINARY; PRT; 142 AA.
 AC 061985;
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DR 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 GN K02H11.6.
 DE Hypothetical 16.2 kDa protein.
 OS Caenorhabditis elegans.
 CC Eukaryota: Metazoa: Nematoda: Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed-9851916;
 RA Waterston R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]

Query Match 2.4%; Score 7; DB 5; Length 142;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 159 KLEIKD 165
 DB 8 KLEIKD 14

RESULT 30

OY 159 KLEIKD 165
 DB 8 KLEIKD 14

Query Match
 Best Local Similarity 2.4%; Score 7; DB 4; Length 148;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 AKLEIKD 164
 DB 31 AKLEIKD 37

RESULT 31

OY 158 AKLEIKD 164
 DB 31 AKLEIKD 37

Query Match
 Best Local Similarity 2.4%; Score 7; DB 4; Length 148;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR EMBL: AK018019; BAB31040.1;
 DR MGI: 1915024; 5830457J20R1K.
 SQ SEQUENCE 150 AA; 16398 MW; 90160AA1E24F1CF CRC64;

Query Match

Best Local Similarity 2.4%; Score 7; DB 11; Length 150;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 AKLEIKD 164
 DB 79 AKLEIKD 85

RESULT 32

OY 158 AKLEIKD 164
 DB 79 AKLEIKD 85

Query Match
 Best Local Similarity 2.4%; Score 7; DB 2; Length 155;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 KATAGDK 33
 DB 92 KATAGDK 98

RESULT 33

OY 27 KATAGDK 33
 DB 92 KATAGDK 98

Query Match 2.4%; Score 7; DB 2; Length 155;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 KAIAGDK 33
 |||||
 DB 92 KAIAGDK 98

RESULT 34
 P91468 PRELIMINARY; PRT; 155 AA.
 ID P91468
 AC P91468;
 DT 01-MAY-1997 (TREMblrel. 03, Created)
 DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE T20D4.7 protein.
 GN T20D4.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Anscough R., Anderson K., Baynes C., Berts M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Lalster N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shomkeen R.,
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Spratt J., Wohldman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans".
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Minx, P., Graves T.;
 RT "The sequence of C. elegans cosmid T20D4.";
 RN Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U80029; AB37590.1; -;
 DR InterPro: IPR000063; ThioRedoxin.
 DR PRINTS: PR00421; THIOREDOXIN.
 SQ SEQUENCE 155 AA; 17629 MW; C2C540B87807DEA CRC64;

Query Match 2.4%; Score 7; DB 5; Length 155;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 KLEKIDK 165
 |||||
 DB 22 KLEKIDK 28

RESULT 35
 Q97KF0 PRELIMINARY; PRT; 155 AA.
 ID Q97KF0
 AC Q97KF0;
 DT 01-OCT-2001 (TREMblrel. 18, Created)
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
 DE Hypothetical protein CAC0969.

GN CAC0969.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Bacilli; Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_Taxid=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.W., Dubois J., Qiu D., Hilti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 bacterium Clostridium acetobutylicum.";
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL: AE007613; AAK78945.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 155 AA; 18691 MW; F684BE8CD10FE26A CRC64;

Query Match 2.4%; Score 7; DB 16; Length 155;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 LEKIDKE 166
 |||||
 DB 32 LEKIDKE 38

RESULT 36
 ID O8VEA3 PRELIMINARY; PRT; 161 AA.
 AC O8VEA3;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Similar to homeo box B2.
 GN A1894218.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC019414; AAH19414.1; -;
 DR MGI: 2144329; A1894218.
 SQ SEQUENCE 161 AA; 16953 MW; F7749E50E542AC37 CRC64;

Query Match 2.4%; Score 7; DB 11; Length 161;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 EPLPEDA 56
 |||||
 DB 93 EPLPEDA 99

RESULT 37
 Q8ZPG2 PRELIMINARY; PRT; 162 AA.
 ID Q8ZPG2
 AC Q8ZPG2;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Putative marR-family transcriptional regulator.
 GN STM1547.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_Taxid=602;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-L72 / SGSC1412 / ATCC 700720;
 RX MEDLINE-21534948; PubMed-11677609;
 RA McClelland M., Sanderson K.E., Spleeh J., Clifton S.W., Latreille P.,
 RA Courtney L., Portwoll S., Ali J., Dente M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 DR EMBL: AE008767; AL20466.1;
 DR InterPro: IPR000485; ASNC_trans_reg.
 DR InterPro: IPR001034; DeOR.
 DR Pfam: PF01047; Mairr. 1.
 DR PRINTS: PR00033; HTHASNC.
 DR SMART: SM00347; HTH_MARR. 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 162 AA; 18280 MW; 043E80C64704A464 CRC64;

Query Match 2.4%; Score 7; DB 16; Length 162;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 VAKLEKL 163
 DB 84 VAKLEKL 90

RESULT 38
 ID 082718 PRELIMINARY; PRT; 162 AA.
 AC 082718;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Putative regulatory protein.
 GN STY1514.
 OS *Salmonella typhi*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Salmonella*.
 OX NCBI_TaxID-601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C718;
 RX MEDLINE-21534947; PubMed-11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebalinda M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant *Salmonella*
 RT *enterica* serovar Typhi C718.";
 RL Nature 413:848-852(2001).
 DR EMBL: AL627270; CAD01769.1;
 DR InterPro: IPR000485; ASNC_trans_reg.
 DR InterPro: IPR000835; HTH_MARR.
 DR Pfam: PF01047; Mairr. 1.
 DR PRINTS: PR00033; HTHASNC.
 DR SMART: SM00347; HTH_MARR. 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 162 AA; 18321 MW; EEC09FC9416872CF CRC64;

Query Match 2.4%; Score 7; DB 16; Length 162;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 157 VAKLEKL 163

DB 84 VAKLEKL 90
 RESULT 39
 ID 092117 PRELIMINARY; PRT; 170 AA.
 AC 092117;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
 DE Putative methylase (Fragment).
 OS *Helicobacter pylori* (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; *Helicobacter* group;
 OX NCBI_TaxID-210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-J166;
 RX MEDLINE-99007275; PubMed-9789049;
 RA Akopyants N.S., Fradkov A., Diachenko L., Hill J.E., Siebert P.D.,
 RA Lukyanov S.A., Sverdlov E.D., Berg D.E.;
 RT "PCR-based subtractive hybridization and differences in gene content
 RT among strains of *Helicobacter pylori*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:13108-13113(1998).
 DR EMBL: AF025981; AAC69266.1;
 DR Methyitransferase.
 FT NON_TER 1 1
 FT NON_TER 170 170
 SQ SEQUENCE 170 AA; 19291 MW; 0879FA8D56A997C CRC64;

Query Match 2.4%; Score 7; DB 2; Length 170;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 AKLEKLD 164
 DB 85 AKLEKLD 91

RESULT 40
 ID 080C67 PRELIMINARY; PRT; 172 AA.
 AC 080C67;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Hypothetical protein Atu2631.
 GN ATU2631 OR AGR_C_4771.
 OS *Agrobacterium tumefaciens* (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID-176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21608550; PubMed-11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D. St.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*
 RT C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21608551; PubMed-11743194;
 Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,

RA Qutrollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Homiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Martelz B.,
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
 RA Cleio C., Slater S.,
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.",
 RL Science 294:2323-2328(2001).
 DR EMBL: AE009210; AAL43612.1; ALT_INIT.
 DR EMBL: AE008176; AAK8352.1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 172 AA; 19276 MW; A0CEA1BE63A16DC5 CRC64;

Query Match 2.4%; Score 7; DB 16; Length 172;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 212 PPDQISS 218
 DB 54 TPDQISS 60

RESULT 41
 ID 016158 PRELIMINARY; PRT; 184 AA.
 AC 016158;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE S2P2 protein (RH5364P).
 GN S2P2 OR CG14904.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERRELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dalke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Hanks M.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui Y., Reese M.G.,
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskys R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodege T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao O., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster.",
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OREGON-R; TISSUE-NEURON;
 RL Kelly L.E., Phillips A.M., Delbridge M., Stewart R.,
 RN Insect Biochem. Mol. Biol. 0:0-0(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OREGON-R; TISSUE-NEURON;
 RA Quattrone A., Xu H., Pacini A., Alkon D.L.,
 RT "Drosophila Cex protein: an interface between Ca2+ oscillations and
 RT cGase activity in neuronal signal transduction.",
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERRELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nuncio J., Pacle J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.,
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003714; AAF53336.1; -
 DR EMBL: AF014952; AAB67805.1; -
 DR EMBL: AF093240; AAC62484.1; -
 DR EMBL: AF089680; AAL90418.1; -
 DR FLYbase: FBgn020907; Scp2.
 DR Interpro: IPR002048; EF-hand.
 DR Pfam: PF00036; EF-hand.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_2.
 DR SEQUENCE 184 AA; 21280 MW; 9C1AE61B6C23810 CRC64;

Query Match 2.4%; Score 7; DB 5; Length 184;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 143 PKNKERY 149
 DB 50 PKNKERY 56

RESULT 42
 ID 0928B7 PRELIMINARY; PRT; 192 AA.
 AC 0928B7;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CT277 SIMILARITY (CPJ0426 protein).
 GN CPN0426 OR CPJ0426 OR CPJ0327.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 CC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OC NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CML029;
 RX MEDLINE=99206606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.,
 RT "Comparative genomics of Chlamydia pneumoniae and C. trachomatis.",
 RN Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heldelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Ufferback T., Berry K., Bass S.,
 RA Liner K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,

```

RA Gwim M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "genome sequences of Chlamydia trachomatis MOPn and Chlamydia
RT Pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shiba T., Ishii K., Hattori M., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CML029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL: AE001625; AAD18570.1;
DR EMBL: AE002194; AAF38182.1;
DR EMBL: AP002546; BAA98634.1;
DR TIGR: CP0327;
RW Complete Proteome.
SQ SEQUENCE 192 AA; 21464 MW; 9C38C329AEDDB76F CRC64;

Query Match 2.4%; Score 7; DB 16; Length 192;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 VAKLEKL 163
Db 126 VAKLEKL 132

RESULT 43
OBYFN3 PRELIMINARY; PRT: 192 AA.
AC OBYFN3;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Colicin V production protein.
GN BMEI1487.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=16M / ATCC 23456 / BIOTYPE 1;
RX MEDLINE=20020109; PubMed=11756688;
RA Delvecchio V.G., Kapriel V., Redkar R.J., Patra G., Muier C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-U.,
RA Haselkorn R., Kyripides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL: AE009585; AAL52668.1;
DR InterPro: IPR003825; Colicin_V.
DR Pfam: PF02674; Colicin_V.1.
KW Complete proteome.
SQ SEQUENCE 192 AA; 20559 MW; D2AD501798E5C04A CRC64;

Query Match 2.4%; Score 7; DB 16; Length 192;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 218 SLTEKLK 224
Db 162 SLTEKLK 168

RESULT 44
O9D920 PRELIMINARY; PRT: 195 AA.
ID O9D920

```

```

AC O9D920;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 583045720Rik protein (RIKEN cDNA 583045J20 gene).
GN 583045720Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Kikawa T., Hara A., Fukunishi Y., Kono H., Aichi S., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hornann M., Humé D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Magarelli J., Kombeerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN [2]
RP SEQUENCE FROM N.A.
RX STRAUSBERG R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK007423; BAB25030.1;
DR EMBL: BC023059; AAR3059.1;
DR MGI: 1915024; 583045720Rik.
SQ SEQUENCE 195 AA; 22120 MW; F3DF9A752D035D51 CRC64;

Query Match 2.4%; Score 7; DB 11; Length 195;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 158 AKLEKLD 164
Db 79 AKLEKLD 85

RESULT 45
O969J3 PRELIMINARY; PRT: 196 AA.
ID O969J3;
AC O969J3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE LOHICR12 (Hypothetical 22.2 kDa protein).
GN LOHICR12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX Montpetit A., Bolly G., Sinnott D.;
RT "A detailed transcriptional map of the chromosome 12p12 tumor
RT suppressor locus.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

RC TISSUE-BRAIN;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY037865; AK71328.1; -
DR EMBL; BC013668; AH13668.1; -
KW Hypothetical protein.
SQ SEQUENCE 196 AA; 22222 MW; CC369C1971B4FEB8 CRC64;

Query Match: 2.4%; Score 7; DB 4; Length 196;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

OY 158 AKLEKID 164
DB 79 AKLEKID 85

Search completed: May 21, 2003, 11:12:54
Job time : 37 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 10:54:37 ; Search time 15 Seconds
(without alignments)
562.959 Million cell updates/sec

Title: US-09-869-677a-2

Perfect score: 1475

Sequence: 1 STGAKTKAKSDKLKVATNS.....PGDSYAMMKWLDKISEGL 287

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep.*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1179.5	80.0	309	2	US-08-715-131-2
2	1179.5	80.0	309	4	US-09-221-753-2
3	1174.5	79.6	289	4	US-08-961-083-20
4	1165.5	79.0	310	1	US-07-791-377-2
5	1165.5	79.0	310	4	US-08-356-106-2
6	1142	77.4	293	4	US-09-071-035-496
7	1142	77.4	316	4	US-09-071-035-494
8	892	60.5	289	4	US-09-071-035-28
9	892	60.5	306	4	US-09-071-035-26
10	748.5	50.7	318	4	US-09-134-001C-5547
11	708	48.0	309	1	US-08-729-202-1
12	708	48.0	309	1	US-08-896-371-1
13	314	21.3	292	2	US-08-928-284-2
14	307.5	20.8	318	1	US-08-599-480-2
15	307.5	20.8	318	2	US-08-842-199-2
16	307.5	20.8	318	5	PCT-US95-00190-2
17	262	17.8	344	4	US-09-134-001C-3524
18	258.5	17.5	197	4	US-08-961-083-108
19	253	17.2	306	4	US-09-160-975A-2
20	224.5	15.2	320	4	US-09-134-001C-4439
21	170	11.5	163	4	US-08-961-083-50
22	142	9.6	337	3	US-08-686-528A-2
23	142	9.6	337	4	US-09-456-287-2
24	141.5	9.6	313	3	US-08-686-528A-3
25	141.5	9.6	313	4	US-09-456-287-3
26	103	7.0	372	1	US-07-813-584A-3
27	103	7.0	372	1	US-08-330-515-3

28	102.5	6.9	489	1	US-08-318-831-4	Sequence 4, Appl1
29	102.5	6.9	666	1	US-08-318-831-3	Sequence 3, Appl1
30	102.5	6.9	814	1	US-08-318-831-2	Sequence 2, Appl1
31	102	6.9	565	4	US-08-961-083-218	Sequence 218, App
32	94.5	6.4	976	4	US-09-104-324B-4	Sequence 4, Appl1
33	94.5	6.4	1128	4	US-08-923-992A-6	Sequence 6, Appl1
34	94.5	6.4	1209	5	PCT-US95-04589-107	Sequence 107, App
35	94.5	6.4	1258	2	US-08-310-912A-107	Sequence 107, App
36	94.5	6.4	1258	4	US-09-301-085-107	Sequence 107, App
37	93.5	6.3	632	4	US-09-016-080-1	Sequence 1, Appl1
38	89.5	6.1	484	2	US-08-879-260-4	Sequence 4, Appl1
39	89.5	6.1	484	3	US-09-231-529-4	Sequence 4, Appl1
40	89.5	6.1	484	4	US-08-977-816-4	Sequence 4, Appl1
41	89.5	6.1	1098	4	US-08-923-992A-8	Sequence 8, Appl1
42	89.5	6.1	1104	4	US-08-923-992A-4	Sequence 4, Appl1
43	89.5	6.1	1164	4	US-08-923-992A-2	Sequence 2, Appl1
44	89.5	6.1	2285	4	US-09-308-375-2	Sequence 2, Appl1
45	89	6.0	437	4	US-08-822-774-51	Sequence 51, Appl1

ALIGNMENTS

RESULT 1
US-08-715-131-2
; Sequence 2, Application US/08715131
; Patent No. 5854416
GENERAL INFORMATION:
APPLICANT: Sampson, Jacquelyn S.
APPLICANT: Russell, Harold
APPLICANT: Tharpe, Jean A.
APPLICANT: Ades, Edwin W.
APPLICANT: Carlone, George M.
TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE 37-KDA SURFACE
TITLE OF INVENTION: ADHESION A PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Needle & Rosenberg, P.C.
STREET: 127 Peachtree Street, Suite 1200
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303-1811
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/715,131
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Spratt, Gwendolyn D.
REGISTRATION NUMBER: 36, 016
REFERENCE/DOCKET NUMBER: 14114.0200
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 688-0770
TELEFAX: (404) 688-9880
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 309 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-715-131-2

Query Match 80.0%; Score 1179.5; DB 2; Length 309;
Best Local Similarity 78.4%; Pred No. 1.7e-99;
Matches 225; Conservative .30; Mismatches 31; Indels 1; Gaps 1;
:::|||||

Db 21 ASGKDTSGOKLVATNSIADITKNIAGDKIDLSIVPIGODPHEPEPLPEDVAKTS 80
QY 61 NADVIFYNGINLEDGGAQAFETKLVAKNAOKTKNDYFAVSDGIDVYILEGASKEGKEDPHA 120
Db 81 EADLIIFYNGINLETGGMNFTLVENAKKTEKDFAVSDGVYILEGONKKGKEDPHA 140
QY 121 WNLNENGIIFYSKNIKOLIAKDPKNEKTEYKNIKAVALEKIDKDEAKSKPDIAENKRL 180
Db 141 WNLNENGIIFYFAKNIKOLIAKDPKNEKTEYKNIKAVALEKIDKDEAKSKPDIAENKRL 200
QY 181 IYTSSEGFYFSKAYGVPASVAYIWEINTEEEGTPDOJSSILEKIKVIPSALFVSSVDNR 240
Db 201 IYTSSEGFYFSKAYGVPASVAYIWEINTEEEGTPDOJSSILEKIKVIPSALFVSSVDNR 260
QY 241 PMETVSKDSGIPYSEIFETDSIAKKGKPDGSYAMKKNLIDKISGL 287
Db 261 PMKTVSODTNIPYIAQIFETDSIAEGKGDSTYSMMKYNLDKIAEGL 307

RESULT 2
US-09-221-753-2
Sequence 2, Application US/09221753
Patent No. 6217884
GENERAL INFORMATION:
APPLICANT: SAMSON, JACQUELYN S.
APPLICANT: RUSSELL, HAROLD
APPLICANT: THARPE, JEAN A.
APPLICANT: ADES, EDWIN W.
APPLICANT: CARLONE, GEORGE M.
TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE 37 kDa SURFACE
FILE REFERENCE: 64778 US
CURRENT APPLICATION NUMBER: US/09/221.753
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: US 07/791,377
EARLIER FILING DATE: 1991-09-17
EARLIER APPLICATION NUMBER: US 07/816,286
EARLIER FILING DATE: 1993-01-03
EARLIER APPLICATION NUMBER: US 08/222,179
EARLIER FILING DATE: 1994-04-04
EARLIER APPLICATION NUMBER: US 08/715,131
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 309
TYPE: PRT
ORGANISM: STREPTOCOCCUS PNEUMONIAE
US-09-221-753-2

Query Match 80.0%; Score 1179.5; DB 4; Length 309;
Best Local Similarity 78.4%; Pred. No. 1.7e-99;
Matches 223; Conservative 30; Mismatches 31; Indels 1; Gaps 1;
QY 2 STGAK-TAKSDKLKVVATNSIADITKNIAGDKIDLSIVPIGODPHEPEPLPEDVAKTS 60
Db 21 ASGKDTSGOKLVATNSIADITKNIAGDKIDLSIVPIGODPHEPEPLPEDVAKTS 80
QY 61 NADVIFYNGINLEDGGAQAFETKLVAKNAOKTKNDYFAVSDGIDVYILEGASKEGKEDPHA 120
Db 81 EADLIIFYNGINLETGGMNFTLVENAKKTEKDFAVSDGVYILEGONKKGKEDPHA 140
QY 121 WNLNENGIIFYSKNIKOLIAKDPKNEKTEYKNIKAVALEKIDKDEAKSKPDIAENKRL 180
Db 141 WNLNENGIIFYFAKNIKOLIAKDPKNEKTEYKNIKAVALEKIDKDEAKSKPDIAENKRL 200
QY 181 IYTSSEGFYFSKAYGVPASVAYIWEINTEEEGTPDOJSSILEKIKVIPSALFVSSVDNR 240
Db 201 IYTSSEGFYFSKAYGVPASVAYIWEINTEEEGTPDOJSSILEKIKVIPSALFVSSVDNR 260
QY 241 PMETVSKDSGIPYSEIFETDSIAKKGKPDGSYAMKKNLIDKISGL 287
Db 261 PMKTVSODTNIPYIAQIFETDSIAEGKGDSTYSMMKYNLDKIAEGL 307

RESULT 3
US-08-961-083-20
Sequence 20, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 Inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-20

Query Match 79.6%; Score 1174.5; DB 4; Length 289;
Best Local Similarity 78.0%; Pred. No. 4.4e-99;
Matches 224; Conservative 31; Mismatches 31; Indels 1; Gaps 1;
QY 2 STGAK-TAKSDKLKVVATNSIADITKNIAGDKIDLSIVPIGODPHEPEPLPEDVAKTS 60
Db 1 ASGKDTSGOKLVATNSIADITKNIAGDKIDLSIVPIGODPHEPEPLPEDVAKTS 60
QY 61 NADVIFYNGINLEDGGAQAFETKLVAKNAOKTKNDYFAVSDGIDVYILEGASKEGKEDPHA 120
Db 61 EADLIIFYNGINLETGGMNFTLVENAKKTEKDFAVSDGVYILEGONKKGKEDPHA 120
QY 121 WNLNENGIIFYSKNIKOLIAKDPKNEKTEYKNIKAVALEKIDKDEAKSKPDIAENKRL 180
Db 121 WNLNENGIIFYFAKNIKOLIAKDPKNEKTEYKNIKAVALEKIDKDEAKSKPDIAENKRL 180
QY 181 IYTSSEGFYFSKAYGVPASVAYIWEINTEEEGTPDOJSSILEKIKVIPSALFVSSVDNR 240
Db 181 IYTSSEGFYFSKAYGVPASVAYIWEINTEEEGTPDOJSSILEKIKVIPSALFVSSVDNR 240
QY 241 PMETVSKDSGIPYSEIFETDSIAKKGKPDGSYAMKKNLIDKISGL 287
Db 241 PMKTVSODTNIPYIAQIFETDSIAEGKGDSTYSMMKYNLDKIAEGL 287

RESULT 4
US-07-791-377-2
Sequence 2, Application US/07791377

```

; Patent No. 5422427
; GENERAL INFORMATION:
; APPLICANT: Russell, Harold
; APPLICANT: Tharpe, Jean A.
; APPLICANT: Sampson, Jacquelyn
; APPLICANT: O'Connor, Steven P.
; TITLE OF INVENTION: PNEUMOCOCCAL FIMBRIAL PROTEIN A
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L. Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/791,377
; FILING DATE: 19911121
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/91969
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-791-377-2

Query Match          79.0%; Score 1165.5; DB 1; Length 310;
Best Local Similarity 76.8%; Pred. No. 3.2e-98;
Matches 219; Conservative 33; Mismatches 32; Indels 1; Gaps 1;

QY 3 TGAKTAKSKLKVATNSIADMTKAIAGDKIDLSIVPGDPHEPEPLPEDAEKTSNA 62
DB 25 TSKSS-SDKLKVTYNTSLADITKNIAGDKIELHSIVPGDPHEPEPLPEDVKKTSQA 83
QY 63 DVIFYNGINLEDGGQAMFTKLVNAQTKNKDYFAVSDGIDVYLLGASGKEDPHAWL 122
DB 84 DLIFYNGINLEDGGQAMFTKLVNAQTKNKDYFAVSDGIDVYLLGASGKEDPHAWL 143
QY 123 NLENGIITSKNIAKOLIAKDPKPKETYEKNLKYAVAKLEKLDKREKSPDAIAENKKLIY 182
DB 144 NLENGIITSKNIAKOLIAKDPKPKETYEKNLKYAVAKLEKLDKREKSPDAIAENKKLIY 203
QY 183 TSGGCRFYFSKAYGVSAIYWEINTEVEGTPEDIKTLERQTKVPSLVESSVDERPM 242
DB 204 TSGGCRFYFSKAYGVSAIYWEINTEVEGTPEDIKTLERQTKVPSLVESSVDERPM 263
QY 243 ETVSKDSGIPYSEIFTDIAKAGKPGDSYAAAMKNLKDISEGL 287
DB 264 KTVSKDSNIPIFAKIFTDSIAKGEEDGSYSMMKNLKEKIAISGL 308

RESULT 5
US-08-356-106-2
; Sequence 2, Application US/08356106
; Patent No. 6312944
; GENERAL INFORMATION:
; APPLICANT: Russell, Harold
; APPLICANT: Tharpe, Jean A.
; APPLICANT: Sampson, Jacquelyn

```

```

; APPLICANT: O'Connor, Steven P.
; TITLE OF INVENTION: PNEUMOCOCCAL FIMBRIAL PROTEIN A
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L. Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,106
; FILING DATE: 15-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/791,377
; FILING DATE: 17-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Scott, Watson T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/91969
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 310 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-106-2

Query Match          79.0%; Score 1165.5; DB 4; Length 310;
Best Local Similarity 76.8%; Pred. No. 3.2e-98;
Matches 219; Conservative 33; Mismatches 32; Indels 1; Gaps 1;

QY 3 TGAKTAKSKLKVATNSIADMTKAIAGDKIDLSIVPGDPHEPEPLPEDAEKTSNA 62
DB 25 TSKSS-SDKLKVTYNTSLADITKNIAGDKIELHSIVPGDPHEPEPLPEDVKKTSQA 83
QY 63 DVIFYNGINLEDGGQAMFTKLVNAQTKNKDYFAVSDGIDVYLLGASGKEDPHAWL 122
DB 84 DLIFYNGINLEDGGQAMFTKLVNAQTKNKDYFAVSDGIDVYLLGASGKEDPHAWL 143
QY 123 NLENGIITSKNIAKOLIAKDPKPKETYEKNLKYAVAKLEKLDKREKSPDAIAENKKLIY 182
DB 144 NLENGIITSKNIAKOLIAKDPKPKETYEKNLKYAVAKLEKLDKREKSPDAIAENKKLIY 203
QY 183 TSGGCRFYFSKAYGVSAIYWEINTEVEGTPEDIKTLERQTKVPSLVESSVDERPM 242
DB 204 TSGGCRFYFSKAYGVSAIYWEINTEVEGTPEDIKTLERQTKVPSLVESSVDERPM 263
QY 243 ETVSKDSGIPYSEIFTDIAKAGKPGDSYAAAMKNLKDISEGL 287
DB 264 KTVSKDSNIPIFAKIFTDSIAKGEEDGSYSMMKNLKEKIAISGL 308

RESULT 6
US-09-071-035-496
; Sequence 496, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gili H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.

```

[illegible]

Query Match	60.5%;	Score 892;	DB 4;	Length 289;
Best Local Similarity	-58.4%;	Pred. No. 2.2e-73;		
Matches 163; Conservative	48;	Mismatches 68;	Indels 0;	Gaps 0;

RESULT 9
US-09-071-035-26

GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

Query Match	60.5%;	Score 892;	DB 4;	Length 308;
Best Local Similarity	58.4%;	Pred. NO. 2.4e-73;		
Matches 163;	Conservative 48;	Mismatches 68;	Indels 0;	Gaps 0;

RESULT 10
US-09-134-001C-5547

Query Match	50.7%;	Score 748.5;	DB 4;	Length 316;
Best Local Similarity	52.9%;	Pred. No. 2.8e-60;		
Matches 145; Conservative	52;	Mismatches 74;	Indels 3;	Gaps 3;

OY 12 K L K V A T N S I I A D M T R A I A G D K I D L H S I P I C O D P H E E P L B D A E K T S N A D V I F Y N G I N 71
| | | | | | | | :
D b 39 K L K V A T N S I I Y D M K R V G N K V D V H S I V P C G D P H E E V K R P P D I K A L T D A D A V F Y N G I N 98

Qy	72	LEDGQAEFTLYVNAOK - TNKQYFVNSODIYILEG - ASEKGRDPAHMLNGLII	129
Db	99	LET- GNGHEKALDQAGKSTKDKVANIASANNKFTLYLNGEGRKNNKODPAHMSLENGIK	157
Qy	130	YSKNIAKOLIAKDRPNKTEYENKLAAYAKLEKLDKEASKFDAIAENKKLIVSECCFK	189
Db	158	YKTKQKSLHHDKKDKSTYEGCNAYISKLEELTNKNNKFFDPKNNRPMYSDEAFK	217
Qy	190	YFSKAYGVPNAVMEINTEEGCTPDQJSSLEKIKLVKTPSALYVESSVDRRPMETVSKDS	249
Db	218	YFAOQFDVPCPYIWEINTEKOGTSGCOMQAIKFPVKDHLKLLHLETSDVKRAMOSLEET	277
Qy	250	GIPYISEFTDIASKKGRPSDYAAKMNWINDKI	283
Db	278	KKDIIEVFTSIGREGTKGSDYTKMMKMSNDITJ	311

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Qy	3	TGKTA---KSDDKLVATNSIIMDMFKAIGDDILDSHYIPGIDQHEPEPPLEDEKXT	59
		: : : : : : : : :	
Db	20	TGGKSSDKSKSGKGLKVATTNSILYDMAKNVGGDNDVHSHYPVQODHEPEYFKRDLKKS	79
Qy	60	SNADYIFYNGINLEDDGQAMFTKLVKNAQKT-KKKDYFAVSDGIDVYILEG-ASEKGED	117
		: : : : : : : : :	
Db	80	TDADYILYNGINLIER-GNGMEFEKALEQGRKSLKVAVSKDVPYLYLMEENSKKOD	138
Qy	118	PHAMVNLNGIYISNNIKOLIAIDPNKKFEYENKRLKAYVAKLEKDKPEAKSKFEDAEN	177
		: : : : : : : : :	
Db	139	PHAMKFEYIGIKYKTIQDTPLDTTKNKIQLMQKGNKTYILQLEKLNDSIDKFNNDLPKE	196
Qy	178	KLIIVTSEGCERYSKAYGVPSPAYIWEINTEEBESTPDQISLIEKLVKIPSPALVESSV	237
		: : : : : : : : :	
Db	199	ORAMITSEGAKEYESKQGIIPGYIWEINTEEKQSTPEOMRAQIFVYKHKHLKHLVETSV	256
Qy	238	DRPRPETSCKDGPPIYIEIPIITDSIAKKGKSGSYAAMKKNLEKXI	283
		: : : : : : : : :	
Db	259	DKAMESLSSEETKNDIPGEVYTDTSIGKRGSTGDSYTKKMAKSNIEYV	304

RESULT 12
 US-088-896-371-1
 Sequence 1, Application US/08896371
 Patent NO. 5801234
 GENERAL INFORMATION:
 APPLICANT: Hodgson, John
 APPLICANT: Burnham, Martin
 TITLE OF INVENTION: NOVEL SALIYA BINDING PR
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESS: SmithKline Beecham Corporation
 STREET: 709 Swedeland Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: U.S.A.
 ZIP: 19406-0939
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/896,371
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/729,202
 FILING DATE:
 APPLICATION NUMBER: 9521147.0
 FILING DATE: 16-OCT-1995
 APPLICATION NUMBER: 9604599.2
 FILING DATE: 04-MAR-1996
 APPLICATION NUMBER: 9616136.9
 FILING DATE: 01-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Gimmli, Edward R
 REGISTRATION NUMBER: 38, 891
 REFERENCE/DOCKET NUMBER: P31279
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090
 TELEX:
 INFORMATION FOR SEQ. ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 309 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHEICAL: NO
 ANTI-SENSE: NO

;
; . FRAGMENT TYPE: N-terminal
; . ORIGINAL SOURCE:
US-08-896-371-1

Query Match	48.0%;	Score 708;	DB 1;	Length 309;
Best Local Similarity	49.0%;	Pred. No. 1.3e-56;		
Matches 140;	Conservative 53;	Mismatches 87;	Indels 6;	Gaps 4;

[illegible]

RESULT 13
US-08-928-284-2
; Sequence 2, Application US/08928284

GENERAL INFORMATION:
APPLICANT: Burnham, Martin K.
TITLE OF INVENTION: No. 5882871e1 Saliva Binding Prot
TITLE OF INVENTION: ein
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

```

1 ADDRESS: Dechert, Price & Rhoads
2 STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
3 CITY: Philadelphia
4 STATE: PA
5 COUNTRY: USA
6 ZIP: 19103-2793
7 COMPUTER READABLE FORM:

```

```

?      MEDIUM TYPE: Diskette
?      COMPUTER: IBM Compatible
?      OPERATING SYSTEM: DOS
?      SOFTWARE: FASTSEQ for Windows Version 2.0
?
CURRENT APPLICATION DATA:
APPLICATION NUMBER:   US/08/928,284
FILING DATE:

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/026,682
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:

NAME: Dickinson, O. Todd
REGISTRATION NUMBER: 28,354
REFERENCE/DOCKET NUMBER: P50543
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215/994-2252
TELEFAX: 215/994-2222
TELEX:

```

: INFORMATION FOR SEQ ID NO: 2:
:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 292 amino acids
:   TYPE: amino acid
:   STRANDEDNESS: double
:   TOPOLOGY: linear
US-08-928-284-2

```

Query Match	21.38;	Score 314;	DB 2;	Length 292;
Best Local Similarity	29.98;	Pred. No. 8.6e-21;		
Matches 85;	Conservative 47;	Mismatches 110;	Indels 42;	Gaps 8;

```

QY      12 KLVVAANSIIADMTKVIAGKIDHSHVPIGDPHEHYEPYEDPAEKTSNADVIFYGIN  71
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      28 KLVITVTFPIYVETTKOYAGDTANVELLIGAGTDPHEHYEPBARKAVAKIODADTVYENEN  87

QY      72 LEDGGAMETKLVNAOKTKNDYFAVSDGIDVITYLEGASER-----GRK-----DPh 119
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      88 ME-----TWVPKLLDTLPDKKKVYTIKATG-----DMLLLPGGEEEGDHGEGGHNEEDPh 140

QY      120 AMLNLENGCIYSKRIANOLLAKDPKNNEYKKNLKAVVAKIEKIDREKASKFOALAE-- 176
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      141 VMLSPVRAIKIVENHIDSLSDADYDPDKKETFEKNNAAATIEKIQSIDK-----AYAGGLS 193

QY      177 ---NKKLIYTSCECKRYFASKAVGSAYIWEIINTEEEGTDPDQISLSLEKLVKIPSAIWE 234
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      194 QAKOKSFYTOHAANNYLALDYLGLQVAAISGLSPAEPSAARLAEITVEYKNNKRAYIYFE 253

QY      235 SSVDRPMETYSKDSGT-----PIYSEIFPDSIAKKKGKPDYS 272
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      254 ENASQALANLITSKENGATDVLANPLESITIEDT-----KAGENT 292

```

RESULT 14
US-08-599-480-2
; Sequence 2, Application US/08599480
; Patent No. 5753459

GENERAL INFORMATION:
APPLICANT: Bianco, David R.
APPLICANT: Miller, James N.
APPLICANT: Lovett, Michael A.
APPLICANT: Champion, Cheryl I.
APPLICANT: Tempst, Paul J.
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF A
TITLE OF INVENTION: T. Pallidum RARE OUTER MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESS: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:

? MEDIUM TYPE: Floppy disk
 ? COMPUTER: IBM PC compatible
 ? OPERATING SYSTEM: PC-DOS/MS-DOS
 ? SOFTWARE: Patentin Release #1.0, Version #1.25
 ?
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/08/599,480
 ? FILING DATE: 23-JAN-1996

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Learn, June M.
REGISTRATION NUMBER: 31,238
REFERENCE/DOCKET NUMBER: 07419/018001 (CIP of 016001)

```

TELECOMMUNICATION INFORMATION
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid

```

```

;          TOPOLOGY:  linear
;          MOLECULE TYPE:  protein
US-08-599-480-2

```

Query Match	20.8%	Score 307.5;	DB 1;	Length 318;
Best Local Similarity	29.0%	Pred. No. 3.8e-20;		
Matches 84; Conservative	57;	Mismatches 136;	Indels 13;	Gaps 4;

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 10:55:06 ; Search time 698 Seconds

(without alignments)
40.777 Million cell updates/sec

Title: US-09-869-677A-2

Perfect score: 1475

Sequence: 1 SSTGAKTKAKSDKLKLVATNS.....PGDSYAMKNNLDKISEGL 287

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PC1_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1174.5	79.6	289	10	US-09-765-272-20 Sequence 20, Appl
2	1174.5	79.6	309	9	US-09-769-787-156 Sequence 156, App
3	260.5	17.7	318	9	US-09-738-626-3532 Sequence 3532, App
4	258.5	17.5	197	10	US-09-765-272-108 Sequence 108, App
5	254	17.2	326	10	US-09-841-132-585 Sequence 585, App
6	228	13.5	327	10	US-09-841-132-515 Sequence 515, App
7	170	11.5	163	10	US-09-765-272-50 Sequence 50, Appl
8	153	10.4	314	9	US-09-738-626-6390 Sequence 6390, App
9	147.5	10.0	278	10	US-09-866-468-17 Sequence 17, Appl
10	120	8.1	27	10	US-09-027-956-7 Sequence 7, Appl
11	102	6.9	565	10	US-09-765-272-218 Sequence 218, App
12	98	6.6	497	9	US-09-820-843A-32 Sequence 32, Appl
13	98	6.6	552	10	US-09-817-764-4 Sequence 4, Appl
14	94.5	6.4	1258	10	US-09-867-852-107 Sequence 107, App
15	94.5	6.4	2150	9	US-10-135-322-17 Sequence 17, Appl
16	94	6.4	1054	9	US-10-217-700-3 Sequence 3, Appl
17	92	6.2	399	9	US-09-769-787-47 Sequence 47, Appl
18	92	6.2	468	9	US-10-077-040-1 Sequence 1, Appl
19	92	6.2	468	9	US-10-063-547-150 Sequence 150, App

20	92	6.2	468	9	US-10-036-041-67 Sequence 67, Appl
21	92	6.2	468	9	US-10-028-072-90 Sequence 90, Appl
22	92	6.2	468	9	US-10-035-855-67 Sequence 67, Appl
23	92	6.2	468	9	US-10-063-616-150 Sequence 150, App
24	92	6.2	468	9	US-10-063-502-150 Sequence 150, App
25	92	6.2	468	9	US-10-121-049-90 Sequence 90, Appl
26	92	6.2	468	9	US-10-123-904-90 Sequence 90, Appl
27	92	6.2	468	9	US-10-140-470-90 Sequence 90, Appl
28	92	6.2	468	9	US-09-931-836-67 Sequence 67, Appl
29	92	6.2	468	9	US-10-173-746-90 Sequence 90, Appl
30	92	6.2	468	9	US-10-176-918-90 Sequence 90, Appl
31	92	6.2	468	9	US-10-176-921-67 Sequence 67, Appl
32	92	6.2	468	9	US-10-036-214-67 Sequence 67, Appl
33	92	6.2	468	9	US-10-137-865-90 Sequence 90, Appl
34	92	6.2	468	9	US-10-140-474-90 Sequence 90, Appl
35	92	6.2	468	9	US-10-035-719-67 Sequence 67, Appl
36	92	6.2	468	9	US-10-142-431-90 Sequence 90, Appl
37	92	6.2	468	9	US-10-143-114-90 Sequence 90, Appl
38	92	6.2	468	9	US-10-140-002-90 Sequence 90, Appl
39	92	6.2	468	9	US-10-036-160-67 Sequence 67, Appl
40	92	6.2	468	9	US-10-142-419-90 Sequence 90, Appl
41	92	6.2	468	9	US-10-035-958-67 Sequence 67, Appl
42	92	6.2	468	9	US-10-036-150-67 Sequence 67, Appl
43	92	6.2	468	9	US-10-063-518-150 Sequence 150, App
44	92	6.2	468	9	US-10-063-598-150 Sequence 150, App
45	92	6.2	468	9	US-10-123-262-90 Sequence 90, Appl

ALIGNMENTS

RESULT 1
US-09-765-272-20
Sequence 20, Application US/09765272
Patent No: US20020061545A1

GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 20:


```

STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 108:
US-09-765-272-108

```

```

Query Match          17.5%; Score 258.5; DB 10; Length 197;
Best Local Similarity 33.2%; Pred. No. 5e-13; Indels 31; Gaps 6;
Matches 67; Conservative 26; Mismatches 78;
QY 12 KLVVATNSIADMTKAIAGDKIDHSIVPIGDDPHEPEPLPEDAKTSNADVIFYNGIN 71
DB 9 KLVVATNSIADMTKAIAGDKIDHSIVPIGDDPHEPEPLPEDAKTSNADVIFYNGIN 68
QY 72 LEDGOAFYTKLVNAQTKNDYPAVSDGIDVYILEGASEK-----GKE-----DPH 119
DB 69 ME---TWPKLIDTLDKRRKYTKATG---DMLLPGGEEEGDHDGEEGHHEFDPH 121
QY 120 AMLNLENGIYSKNIKQIADPKNKETYEKNKAYAKLEKIDPEAKSKRDATAE--- 176
DB 122 VWLSVPAIKLVHEHPRHSADYPKKETFENKAAVYIEKLDALDK-----ATAEGLS 174
QY 177 --NKKLIYTSSECFYFSKAYG 196
DB 175 QAKRSFVTOHAAFVYLDVYG 196

```

```

RESULT 5
US-09-841-132-585
Sequence 585, Application US/09841132
Patent No. US20020061848A1
GENERAL INFORMATION:
APPLICANT: Bhatia, Ajay
APPLICANT: Skeiky, Yasir A.W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
FILE REFERENCE: 210121.469C8
CURRENT APPLICATION NUMBER: US/09/841,132
CURRENT FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 599
SOFTWARE: FastSeq for Windows Version 3.0/4.0
SEQ ID NO 585
LENGTH: 326
TYPE: PRT
ORGANISM: C. Trachomatis D serovar
US-09-841-132-585

```

```

Query Match          17.2%; Score 254; DB 10; Length 326;
Best Local Similarity 25.2%; Pred. No. 2.1e-12;
Matches 79; Conservative 68; Mismatches 117; Indels 50; Gaps 11;
QY 1 SGTGAKTAKSDKIKVATNSIADMTKAIAGDKIDHSIVPIGDDPHEPEPLPEDAKTS 60
DB 30 SSRGNQPA-DESIYVLSNMRIQDCVSRITGBRVKNIVLIDGADIDPHSEVMKGDDBRA 88
QY 61 NADVIFYNGINLE-----DGGQAFYTKLVNAQTKNDYPAVSDGIDVYILEGAS 111
DB 89 MSQILFCNGIGLEHSASLRKHLEGN---PKVVDLQGRILNKCF-----DL-----S 133
QY 112 EKKGEDPHAMLNLENGIYSKNIKQIADPKNKETYEKNKAYAKLEKIDPEAKSKF 171
DB 134 EECFPDPHITWTDKRWGAIVKEMAAALIQFQYEDDPQKADQLISEKEEDRMAARSL 193
QY 172 DAIAENKRLIYTSSECFYFSKAYGVPSAY---IWE-----INTEEGTPDQISL 219
DB 194 STPEKNRYLVTOHNAFYSYFTRRYLSSDAERVSCEWRSRCISPEGLSPEAQLSIRDMYV 253
QY 220 IEKLVKIRKSALEFVSESVDR---RPMETYSKDSGIRI---YSRIFTDSIAKKGPDSTY 273
DB 254 VEYISANDVEYVFLIEDTLQDALRKIVSCSK-SGQKIRLAKSPLYSDNVC-----DNVF 306
QY 274 AMMKRNLDKISEGL 287
DB 307 STFOHNVRTITEEL 320

```

```

RESULT 6
US-09-841-132-515
Sequence 515, Application US/09841132
Patent No. US20020061848A1
GENERAL INFORMATION:
APPLICANT: Bhatia, Ajay
APPLICANT: Skeiky, Yasir A.W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
FILE REFERENCE: 210121.469C8
CURRENT APPLICATION NUMBER: US/09/841,132
CURRENT FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 599
SOFTWARE: FastSeq for Windows Version 3.0/4.0
SEQ ID NO 515
LENGTH: 327
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-841-132-515

```

```

Query Match          15.5%; Score 228; DB 10; Length 327;
Best Local Similarity 24.1%; Pred. No. 2.4e-10;
Matches 77; Conservative 63; Mismatches 119; Indels 60; Gaps 11;
QY 1 SGTGAKTAKSDKIKVATNSIADMTKAIAGDKIDHSIVPIGDDPHEPEPLPEDAKTS 60
DB 28 TNSGFONANS-RPCILSNMRMTHDCYERYVGNRLAVILGSLDPHAEVMAKGDQDKIA 86
QY 61 NADVIFYNGINLEDDGOAFYTKLVNAQTKNDYPAVSDGIDVYILEGA---SEKRE 116
DB 87 GSAVIRFCNGIGLEH-----TLSLRKHLNNPN---SVKLG-ERLIARCAFVLEDDGIC 136
QY 117 DPHAMNL---ENGIIYSKNIKQIADPKNKETYEKNKAYAVALEKIDPEAKSKFPA 173
DB 137 DPHIMWDLISWKEAVI---EITEVLEKEPPEMSAEKASEELVCMSTLDSWAKQCLST 193
QY 174 IAENKRLIYTSSECFYFSKAYGVPSAYIWEINTEEGTPDQISLIEKLVKIPGALFY 233
DB 194 IPENLRIYVSGHNAFYSYFTRRY-----LATPEVASGAMSRCSPEGLSP 239
QY 234 ESSVDR-----PMEYVSKDSGIPYSEIFTSIDIAKKG-KP----- 268

```

Db 240 EAQISVRDIMAAYVYDINEHDSVFEEDTLNODALKIIVSILKSHVLAQKPLYSNDV 299
OY 269 GDSYAMKMNLDKISEL 287
Db 300 DNVSTFKHNVCLTEEL 318

RESULT 7
US-09-765-272-50
; Sequence 50, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P9340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 163 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-09-765-272-50

Query Match 11.5%; Score 170; DB 10; Length 163;
Best Local Similarity 31.1%; Pred. No. 3.6e-06;
Matches 52; Conservative 28; Mismatches 71; Indels 16; Gaps 5;

OY 4 GAKTAKSDK-LKVVATSIADMTKAIAGDKIDLSIYPIGODHEHEPLPEDEKESNA 62
Db 1 GAKESOTGKMKITSTPYPIYAMVKEVSGDLNDR-MIQSSSGIHSEPSANDIAIYDA 59
OY 63 DVIFYNGINLEDGQAWFTKLVKNAQTKRNDYFAVSDGI-----DVIYLEGASEK 113
Db 60 DVIFYHSHTE---SWAGSLDPLRLKSKYK-VLEASEGHTLEVPGLVEDYAGDGVDEK 114
OY 114 GKEDPHAMLENGLIYSKNIAMQLAKDPKNNETYEKRLKAYAVAKL 160
Db 115 TLVDPHTMLDPERKAGEBAQIADKLSEVDSEHKETEQKNQPLSKKL 161

RESULT 8
US-09-738-626-6390
; Sequence 6390, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6390
LENGTH: 314
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6390

Query Match 10.4%; Score 153; DB 9; Length 314;
Best Local Similarity 20.9%; Pred. No. 0.00018;
Matches 68; Conservative 47; Mismatches 101; Indels 110; Gaps 14;

OY 1 SSTAKTKASDKLVYATNSIIADMTKAIAGDKIDLSIYPIG-DHEHEPLPEDEK 59
Db 25 SADSTNAGSLSLKVYSTQVMAVAAEAVPD-VDEAIIIGGIDHSEPSKTDAAV 83
OY 60 SNADVIFYNGINLEDGQAWFTKLVKNAQTKRNDYFAVSDGIDVIYLEGASEK 115
Db 84 SEADIIIVG---GGYDSWL-----YGLEDDDRITIALDSEHDSEHD 125
OY 116 -----BDPHAMLENGLIYSKNIAMQLAKDPKNNETYEK 152
Db 126 DHEHEBAEHEHDERGCHDHDVNEHWYSTE---YSEVAEE----- 166
OY 153 LKAVYANLEKLDKREKSKFDIAENKRL-----IYTSB---GCKYPSKA 194
Db 167 ---FAKVTELDPEAQA---DATAVTTKMDLHNGIDLPVARIQAOTPIADHILSHD 221
OY 195 YGVSAIYIWEINTEBEGTPOQISSLIEKLR-----VIRP-SALFYESSVDRRPMETVS 246
Db 222 ESPEGEYRAVTLSESEPTADVASFODAINNGDLVILYNQSASTVATS-----LKDLA 276
OY 247 KDSGIPY-----SEIFPDSIAK 264
Db 277 EEKGIPVEIETPTONTENFLDAFTK 302

RESULT 9
US-09-886-468-17
; Sequence 17, Application US/09886468
; Patent No. US20020037293A1
; GENERAL INFORMATION:
; APPLICANT: Aventis Pasteur Limited
; TITLE OF INVENTION: Chlamydia antigens and corresponding DNA fragments and uses th
; FILE REFERENCE: 77813-5
; CURRENT APPLICATION NUMBER: US/09/886,468
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/113,280
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113,281
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113,282
; PRIOR FILING DATE: 1998-12-23

```

; PRIOR APPLICATION NUMBER: 60/113,283
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113,284
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113,285
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113,385
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/114,050
; PRIOR FILING DATE: 1998-12-28
; PRIOR APPLICATION NUMBER: 60/114,056
; PRIOR FILING DATE: 1998-12-28
; PRIOR APPLICATION NUMBER: 60/114,057
; PRIOR FILING DATE: 1998-12-28
; PRIOR APPLICATION NUMBER: 60/114,058
; PRIOR FILING DATE: 1998-12-28
; PRIOR APPLICATION NUMBER: 60/114,059
; PRIOR FILING DATE: 1998-12-28
; PRIOR APPLICATION NUMBER: 60/114,061
; PRIOR FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 17
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-869-468-17

```

```

Query Match      10.0%; Score 147.5; DB 10; Length 278;
Best Local Similarity 24.0%; Pred. No. 0.00042;
Matches 63; Conservative 49; Mismatches 98; Indels 53; Gaps 10;

```

```

QY 11 DKLAVATNSIIAMTKAIGDKIDLSHIVPIGDPHEHELEPDAEKTSNADYIFNGI 70
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 24 DKPVILVSIAYKFLVEIAEETCFYATVNMAYDPHTLEPPOQIKELRGDLMFRIG- 82
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 71 NLEDGGAMFTKLVKNQKTKNDYFVSDIDVYLEGASEKRE-----DPRHW 121
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 83 --EAFG-----KNLEKPYMQOVDSQNVSLI-----QCKPCNQHTTVNDFTTW 125
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 122 LNLENGIIYSKNIKQLIADPKRKEITYEKLVAKLEKLDK-----AKSFDAIAE 176
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 126 LSPNKLNVQVETIYTTLSKRPQAHATLYQSNGEKLLALDQLEIILITTSKAK----- 179
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 177 NKKLIIVSEGCERYKAYGVSAYIWEINTEEGTPOJISL---IEKLVIKRSALFV 233
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 180 -QRHILVSHGFYFCFDYFN-SQHTIEKSHVPSPKDVARFRDIDQYKI--SSVILL 235
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 234 ESS-----VDRRPMEYVSKD 248
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 236 EYSGRRSSAMLDARFHHMTVLD 258
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 10
US-09-027-956-7
; Sequence 7, Application US/09027956A
; Patent No. US20010048929A1
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: LINDBERG, Alf
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: NOVEL MULTI-OLIGOSACCHARIDE GLYCOCONTUGATE BACTERIAL
; TITLE OF INVENTION: MENINGITIS VACCINES
; FILE REFERENCE: 1038-791 MTS:1b
; CURRENT APPLICATION NUMBER: US/09/027,956A
; CURRENT FILING DATE: 1998-02-23
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 7
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-027-956-7

```

```

Query Match      8.1%; Score 120; DB 10; Length 27;
Best Local Similarity 88.5%; Pred. No. 0.0031;
Matches 23; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 127 GIYSKNIKQLIADPKRKEITYEK 152
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 1 GIYAKNIKQLIADPKRKNKDEYK 26
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 11
US-09-765-272-218
; Sequence 218, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>

```

```

ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders

```

```

REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:

```

```

TELEPHONE: (301) 309-8512
TELEFAX: (301) 309-8512

```

```

INFORMATION FOR SEQ ID NO: 218:
SEQUENCE CHARACTERISTICS:

```

```

LENGTH: 565 amino acids
TYPE: amino acid

```

```

STRANDEDNESS: single
TOPOLOGY: linear

```

```

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 218:
US-09-765-272-218

```

```

Query Match      6.9%; Score 102; DB 10; Length 565;
Best Local Similarity 23.6%; Pred. No. 3.9;
Matches 61; Conservative 41; Mismatches 97; Indels 60; Gaps 12;

```

```

QY 11 DKLAVATNSIIADMTKAIAGDKIDLSHIVPIGDPHEE-----PLEPDAKTSNA 62
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 231 DKIKIGITKE--PYDKSELNNQIDKASSV-----SPTDYSTASYNALGPVLETAGVYAS 283
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 63 DVIFPENGINLEDGGQAMFTKLVKNQKTKNDYFVSDIDVYLEGASEKGEKDPH--- 119
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 284 EVKQPEVNSE-----TNKLTATIDALNDVKTLELNTI-----ADAKTKVKEHYS 329
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 120 -AMNLENGIIYSKNIKQLIADPKRKEITYEKLVAKLVAKL-----EK----- 162
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 330 RSMQNLQTEVTKAEKVAANTDAKQSEVNEAVEK-LTATIEKLVSEKPIILTFSTDKRI 388
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 163 LDKFAKSKFDALAEKKLI--VTSGCEKRYSKAVGVSAYIWEINTEEGTPOJISLI 220
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 389 LEREAVAKYTLNOMKTKIKSITAE-----LKKGEVINTVVL---TDDKVTETISAA 440
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

OY 221 EKLKVKPSALFEVSSVDR 239
 DB 441 KMLEYKKEVTLSTMTYDR 459

RESULT 12
 US-09-820-843A-32
 ; Sequence 32, Application US/09820843A
 ; Publication No. US20030039963A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Council of Scientific and Industrial Research
 ; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
 ; FILE REFERENCE: 063915
 ; CURRENT APPLICATION NUMBER: US/09/820, 843A
 ; CURRENT FILING DATE: 2001-03-30
 ; NUMBER OF SEQ ID NOS: 118
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 32
 ; LENGTH: 497
 ; TYPE: PRT
 ; ORGANISM: B. burgdorferi
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: predicted coding region BB0553
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: "Xaa" may be any amino acid
 ; OTHER INFORMATION: g1/2688482
 ; US-09-820-843A-32

Query Match
 Best Local Similarity 21.9%; Score 98; DB 9; Length 497;
 Matches 73; Conservative 59; Mismatches 112; Indels 90; Gaps 19;

OY 6 KTAASDKLV-----VATNSI---TADMTKAIAG-----DKIDLSIIV 41
 DB 121 KTSNENKKEISIKKKKKKEITLNKLNKEIVETIKLNKKIKKEDENYENIENIEE 180
 OY 42 TGDOPHE--YEPLEDAEKTSMADVIFNGI--NLEDGQAMFTKLVNAQTKRKNKYFA 97
 DB 181 ETDDDEFDNE-YNDEITEXINEDNYPSNEGIIINLKE-----NLNENETYYA 226
 OY 98 VSGIDIVYEGASEKKEPHAMLNENGIYSKNTAKOLIAKDPKNKETEKN----- 152
 DB 227 INERK---KIDELDRINENENTLIDQREI---RNFKKK--DNSDKNLEIEIENLSSIG 277
 OY 153 -----LKAAYAKLEKLDKEAKSKF-----DAIAENKLIIVSEGCFFYFKAYG 196
 DB 278 RIINDLKRRISANEAIKKQKKIRIDKHLKLEDEKIKENETITLQKELNNFK--- 334
 OY 197 VPSAYIWEITEEGTPDOIS---SLIE--KLK-----VYKPSALFEVSSVDRPPEYV 245
 DB 335 -KEIYQPLN-EETFPSTISKNDLLEENKKLKREYIAKPEKKESRLEENSTPTPTM 392
 OY 246 SKDSGPIYSEITDTSIAKKGKPGDSYAMAKNN 279
 DB 393 IKTADFOIYDILNMYKFKK--GDQ--FAFKKEN 424

RESULT 13
 US-09-817-764-4
 ; Sequence 4, Application US/09817764
 ; Patent No. US20020102678A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HAROCHE, JULIEN
 ; APPLICANT: ALLIGNET, JEANINE
 ; APPLICANT: EL SOLH, NEVINE
 ; TITLE OF INVENTION: STAPHYLOCOCCAL GENE, VGAC, CONFERRING RESISTANCE TO
 ; TITLE OF INVENTION: STREPTOGRAMIN A AND RELATED COMPOUNDS
 ; FILE REFERENCE: 03495, 0201
 ; CURRENT APPLICATION NUMBER: US/09/817,764

;; CURRENT FILING DATE: 2001-03-27
 ;; PRIOR APPLICATION NUMBER: 60/197,372
 ;; PRIOR FILING DATE: 2000-04-14
 ;; NUMBER OF SEQ ID NOS: 11
 ;; SOFTWARE: PatentIn Ver. 2.1
 ;; SEQ ID NO 4
 ;; LENGTH: 552
 ;; TYPE: PRT
 ;; ORGANISM: Staphylococcus sp.
 ; US-09-817-764-4

Query Match
 Best Local Similarity 27.2%; Score 98; DB 10; Length 552;
 Matches 52; Conservative 28; Mismatches 53; Indels 58; Gaps 12;

OY 57 EKTSMADVIFNGINLEDGQAMFTKLVNAQTKRKNRYAVSDGIDVYILEGASEKKE 116
 DB 125 EPTSNLDI---EGI-----ELITWFKFYRDTFLVSH--DRIFLDVCTRIFE 168
 OY 117 DPHAMLNENGI---YSKNI-AKOLIAKDPKNKETEKNKAKVAKLEKDEAKS 169
 DB 169 -----IENGYNIEFGNTNTNTEOKEMILR--KQOEYK---YNSRKOLEQAIRL 215
 OY 170 KPDIAENKRLIYSEGCFFYFKAYGVPASAYIWEI-----NTEEGTPDOISS 218
 DB 216 K-----ENK-----AQGMKPPSKTMTGTSRIWKQKHATKQKKMHNRTKSLTRIDKLN 265
 OY 219 LIEKLVIKPS 229
 DB 266 HVEKIKEL-PS 275

RESULT 14
 US-09-867-852-107
 ; Sequence 107, Application US/09867852
 ; Patent No. US20020147324A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ausubel, Frederick M.
 ; APPLICANT: Staskawicz, Brian J.
 ; APPLICANT: Brent, Andrew F.
 ; APPLICANT: Dahlbeck, Douglas
 ; APPLICANT: Katagiri, Rumiaki
 ; APPLICANT: Kunkel, Barbara N.
 ; APPLICANT: Mindlinos, Michael N.
 ; APPLICANT: Yu, Guo-Liang
 ; TITLE OF INVENTION: RPS2 GENE FAMILY, PRIMERS, PROBES, AND
 ; TITLE OF INVENTION: DETECTION METHODS
 ; FILE REFERENCE: 00786/254002
 ; CURRENT APPLICATION NUMBER: US/09/867,852
 ; CURRENT FILING DATE: 2001-05-29
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/301,085
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-28
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/310,912
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1994-09-22
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 08/227,360
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1994-04-13
 ; NUMBER OF SEQ ID NOS: 208
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 107
 ; LENGTH: 1258
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-867-852-107

Query Match
 Best Local Similarity 19.9%; Score 94.5; DB 10; Length 1258;
 Matches 59; Conservative 51; Mismatches 115; Indels 71; Gaps 10;

OY 4 GAKTASDKLVVATNSIADMTKAIAGDKIDLSIIVPGDPH---EYEPLEDAEKTSS 60
 DB 201 GWHIGKNDKOGAIA-DKVADADWHSKKNLILBDELVDGIDDTHTAVLEKSLDSSENVY 259
 OY 61 MADVIFNGINLEDGQAMFTKL-----VNAQTKRKNKRYAVSDGIDVYILEGAS 111

```
Db 260 MWGLGMGIGKTTAKAVYNNKISSCFDCCCFIDNIRETOEK-----DGVVVLQKLVLS 313
QY 112 EKGKEDPHAMNLNENGIYYSKNIQAOLAKDPKNKETEYKNIKAYVAKLEKLDKEAKSKF 171
Db 314 E-----LIRIDSGSVGFNN-----DSGGRKTIKERYSRFKLVLDVDVEKFKF 357
QY 172 DAIAENKLIYSEGCFFYFSKAYGVPSAYIWEINTEEGTPDQISLIEKLVKIPSA 231
Db 358 EDMGSPKDFI-SQSRFITTSRMV-----LGLINENO-----CKLYEVSMSKPRSL 405
QY 232 FVESSVDRRPMETVSKDGIPIYSEIFTDISLAKGKPGDSYAMMKWNLDKISEGL 287
Db 406 -----ELFSKAFKKNTPSSYETLANDVVDITAGL 437
```

```
RESULT 15
US-10-135-322-17
; Sequence 17, Application US/10135322
; Patent No. US20020173017A1
; GENERAL INFORMATION:
; APPLICANT: BENFEY, PN
; APPLICANT: HEIARIUTTA, Y
; APPLICANT: MAHONEN, AP
; APPLICANT: BONKE, AMM
; APPLICANT: KAUPIINEN, L
; APPLICANT: RIIKONEN, M
; TITLE OF INVENTION: WOODEN LEG GENE, PROMOTER AND USES THEREOF
; FILE REFERENCE: 5914-086-999
; CURRENT APPLICATION NUMBER: US/10/135,322
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 60/253,739
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent version 3.0
; SEQ ID NO 17
; LENGTH: 2150
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-135-322-17
```

Query Match 6.48; Score 94.5; DB 9; Length 2150;
Best Local Similarity 17.7%; Pred. No. 82;
Matches 59; Conservative 52; Mismatches 111; Indels 111; Gaps 13;

```
QY 21 IADWTKAIAGDKIDLH-----STVPIGQ-----DPHEVEPLPEDAKTSNADY 64
Db 1007 VSASIQRLTEBDSLDLYFEDLDSTPIGELIYYRASNAAGNDGSSPTNINMGKLEDRSDM 1066
QY 65 IFYNGINLEDDGQAMFTKLIVNAOKTKNKDY---FAVSDGIDVYILEGASEKGEDPHA 120
Db 1067 IYNTNMV--GGRNMNIAL-RSRKFTNKHFTFYPAIG-GVCMLSALYS-----F 1114
QY 121 WNLNENGIYYSKNIQAOLAKDPKNKETEYKNIKAYVAKLEKLDKEAKSKFDAIAENK 178
Db 1115 WF-----AVNTKHNIKLSATNEDLKEIYNRKLAEKALAESQERLEAMEGSEDAY--- 1165
QY 179 KLIYSEGCFFYFSKAYGVPSAYIWEINTEE-----EGTPDQISLIEKLVK 225
Db 1166 -----WDMKVNITGELHISSRNFQIIKAHDTISQSRITLYEELKS 1203
QY 226 IKPS-----ALFVSSVDRRPMETVS-----KDSGIP 252
Db 1204 SSTNMLNFKGDSKNGSNNGTFNLEKNGKVDSSPSQITNNTTNGGGGELARKNSGYL 1263
QY 253 IYSEIFTDISLAKGKPGDSYAMMKWNLDKISE 285
Db 1264 YNDELFSPIILLEMYVSSPNTHQLAIWKKFLAE 1296
```

Search completed: May 21, 2003, 11:11:10
Job time : 700 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 10:54:37 ; Search time 45 Seconds
(without alignments)
613.124 Million cell updates/sec

Title: US-09-869-677A-2
Pagefoot: 1475

Perfect score: 1475
Sequence: 1 SSTGAKTAKSDKLKVAATNS.....PGDSYYAMKKNLDKISEGL 2877

Scoring table:	BLOSUM62	Canext	0.5
	Canext	10.0	0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%

Listing first 45 summaries

```
Database :
PIR_73:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*
```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1475	100.0	289	2	T48894	lipoprotein mtsA,
2	1179.5	80.0	309	2	E98058	hypothetical protein
3	1179	79.9	309	2	A43583	adhesin B precursor
4	1178	79.9	309	2	A37186	fimbrial adhesin f
5	1174.5	79.6	309	2	H95191	hypothetical protein
6	1165	79.0	310	2	T11551	adhesin - Streptoc
7	1079.5	73.2	313	2	D86789	hypothetical protein
8	804	54.5	310	2	AG1677	adhesin binding p
9	789	53.5	310	2	AG1305	adhesion binding p
10	743	50.4	309	2	H9832	hypothetical prote
11	550	37.3	300	2	D83714	surface adhesin A
12	489.5	33.2	293	2	G64063	probable Mn transp
13	479.5	32.5	333	2	AH0297	periplasmic-bindin
14	475.5	32.5	299	2	AC3106	hypothetical protei
15	475.5	32.2	300	2	A96181	sita protein (Af12
16	457.5	31.0	330	2	S75057	Mn transport protei
17	449.5	30.5	305	2	AI0847	Iron transport prote
18	415.5	28.2	337	2	AD2947	ABC transporter, s
19	415.5	28.2	337	2	F98335	adhesin, probable
20	410.5	27.8	325	2	AI2352	adhesin precursor
21	368.5	25.0	304	2	T45469	cell wall lipoprote
22	348	23.6	319	2	D81182	adhesin, probable
23	344	23.5	319	2	A69756	adhesin protein h
24	342.5	23.2	280	2	C86889	zinc ABC transport
25	341	23.1	308	2	E81923	probable periplasm
26	337	22.8	298	2	E97096	zn-binding lipoprot
27	330	22.4	403	2	D98118	hypothetical protei
28	329	22.3	501	2	T46756	zn-binding lipoprot
29	329	22.3	501	2	F95253	zinc ABC transport

After Filing date
Application

30	328.5	22.3	313	2	AB1094	a probable high-aff
31	326	22.1	312	2	AB1456	probable high-aff
32	325	22.0	306	2	B63992	ABC transporter (m
33	317	21.5	316	2	D71375	probable ABC trans
34	313.5	21.3	326	2	AF2024	hypothetical protea
35	307.5	20.8	308	2	A71360	probable ABC trans
36	304.5	20.6	305	2	F95150	adhesion lipoprote
37	304.5	20.6	311	2	B97985	lipoprotein (lipote
38	301.5	20.4	317	2	AG1283	ABC transporter an
39	289.5	19.6	317	2	AB1655	ABC transporter an
40	281.5	19.1	339	2	AG1910	periplasmic solute
41	271.5	18.4	515	2	G90041	hypothetical prote
42	265.5	18.0	317	2	E83344	probable adhesion
43	254	17.2	336	2	E71561	probable solute pr
44	253	17.2	306	2	T46757	lipoprotein lmb [v
45	251	17.0	332	2	C83682	hypothetical prote

ALIGNMENTS

RESULT 1
T48894
lipoprotein mtsA, metal binding [validated] - Streptococcus pyogenes (Fragment)
N:Alternate names: metal-binding protein mtsA
C:Species: Streptococcus pyogenes
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 2 02-Sep-2000

R; Janulczyk, R.W.; Pallon, D.; Björck, L.

A;Title: Identification and
A:Reference number: 724991

A; Accession: T48894
A; Status: preliminary

A:ResIdues: 1-289 <JAN>

A;Cross-references: EMBL:AF160321
A;Experimental source: strain AP1

A;Gene: mtSA

```
C;Function:      min condensing oac or muen (xrefprocen// modd (max namang p
```

A; Note: protein has affinity for Zn^{2+} , Fe^{2+} and Cu^{2+} ; Zn^{2+} and Cu^{2+} compete

[illegible]

```
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	SSTGAKTKAKSRKLVKAVANSIIADMTKAIADKIDILHSIVPIGDPHFEPLPEDAEKTS	60
Db	1	SSTGAKTKAKSRKLVKAVANSIIADMTKAIADKIDILHSIVPIGDPHFEPLPEDAEKTS	60
QY	61	NADYIFYGINLEBGGQAMFTKLYKNAOKTKNKDYFAVSDGIDVYYLEGASKEKEDPHA	120
Db	61	NADYIFYGINLEBGGQAMFTKLYKNAOKTKNKDYFAVSDGIDVYYLEGASKEKEDPHA	120
QY	121	WLNLENGIITYSKNIAKOLIAADPKNKTEFEKNILKAYAVAKLEKLDKEAKSKDLAENKTL	180
Db	121	WLNLENGIITYSKNIAKOLIAADPKNKTEFEKNILKAYAVAKLEKLDKEAKSKDLAENKTL	180
QY	181	IYTSBGCRRYYSKAYGVASATYWEINTEEBEGTPDOISLIEKLKYIKSALFVSSVDRR	240
Db	181	IYTSBGCRRYYSKAYGVASATYWEINTEEBEGTPDOISLIEKLKYIKSALFVSSVDRR	240
QY	241	PMETYSKDSGIPITYSEIFTDISIAKKGGPGDSYAMAKMNLDKISBGL	287
Db	241	PMETYSKDSGIPITYSEIFTDISIAKKGGPGDSYAMAKMNLDKISBGL	287
RESULT	2		
	598058		

hypothetical protein paa (imported) - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
 C:Accession: E98058
 R:Hoshino, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; H
 e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; N
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: E98058
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-309 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AA00298.1; PID:g15459154; GSPDB:GN00174
 C:Genetics:
 A:Gene: paa
 C:Superfamily: adhesin B

Query Match 80.0%; Score 1179.5; DB 2; Length 309;
 Best Local Similarity 78.4%; Pred. No. 3.6e-72;
 Matches 225; Conservative 30; Mismatches 31; Indels 1; Gaps 1;

OY 2 SNGAK-TAKSDKLKVVATSIADMTAKAGDKIDLSIYVIGODPHEPEPLPDAKTS 60
 DB 21 ASGKKTGSGQKLVVATSIADITKNTAGDKIDLSIYVIGODPHEPEPLPDAKTS 80
 OY 61 NADVIFYNGINLEGGQAMFTKLVNAOKTKKNDYFAVSDGIDVYILEGASKEGDEPNA 120
 DB 81 EADLIFYNGINLEGGQAMFTKLVNAOKTKKNDYFAVSDGIDVYILEGASKEGDEPNA 140
 OY 121 WNLNGLIYSKNIKAKOLIAKPPKNEYKLVKAVAKLEKIDKESKDEKIPAEKKL 180
 DB 141 WNLNGLIYSKNIKAKOLIAKPPKNEYKLVKAVAKLEKIDKESKDEKIPAEKKL 200
 OY 181 IYTSCECFYFSGKAVGSAYIWEINTEEGPDDISSLEKLYKISALFVSSVDR 240
 DB 201 IYTSCECFYFSGKAVGSAYIWEINTEEGPDDISSLEKLYKISALFVSSVDR 260
 OY 241 PMETVSKDSGIPYSEIFTDISIAKKGPDSDYAMKNNLKDISEGL 287
 DB 261 PMETVSKDSGIPYSEIFTDISIAKKGPDSDYAMKNNLKDISEGL 307

RESULT 3

A43583
 Adhesin B precursor - Streptococcus sanguis
 C:Species: Streptococcus sanguis
 C:Date: 12-Jan-1993 #sequence_revision 12-Jan-1993 #text_change 24-Sep-1999
 C:Accession: A43583
 R:Ganeskhumar, N.; Hannam, P.M.; Kolenbrander, P.E.; McBride, B.C.
 Infect. Immun. 59, 1093-1099, 1991
 A:Title: Nucleotide sequence of a gene coding for a saliva-binding protein (saab) from S
 A:Reference number: A43583; MUID:91147187; PMID:16711775
 A:Accession: A43583
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-309 <GAN>
 A:Cross-references: GB:M63481; NID:g153825; PIDN:AA098426.1; PID:g153826
 C:Superfamily: adhesin B

Query Match 79.9%; Score 1179; DB 2; Length 309;
 Best Local Similarity 79.6%; Pred. No. 3.6e-72;
 Matches 226; Conservative 26; Mismatches 30; Indels 2; Gaps 1;

OY 6 KTAQSD--KLVVATSIADMTAKAGDKIDLSIYVIGODPHEPEPLPDAKTSNAD 63
 DB 24 KTAQSD--KLVVATSIADMTAKAGDKIDLSIYVIGODPHEPEPLPDAKTSNAD 83
 OY 64 VIFPNGINLEGGQAMFTKLVNAOKTKKNDYFAVSDGIDVYILEGASKEGDEPNA 123
 DB 84 VIFPNGINLEGGQAMFTKLVNAOKTKKNDYFAVSDGIDVYILEGASKEGDEPNA 143

OY 124 LENGITYSKNIKAKOLIAKPPKNEYKLVKAVAKLEKIDKESKDEKIPAEKKL 183
 DB 144 LENGITYSKNIKAKOLIAKPPKNEYKLVKAVAKLEKIDKESKDEKIPAEKKL 203
 OY 184 SSGCKRYFSGKAVGSAYIWEINTEEGPDDISSLEKLYKISALFVSSVDRPME 243
 DB 204 SSGCKRYFSGKAVGSAYIWEINTEEGPDDISSLEKLYKISALFVSSVDRPME 263
 OY 244 TVSKDSGIPYSEIFTDISIAKKGPDSDYAMKNNLKDISEGL 287
 DB 264 TVSKDSGIPYSEIFTDISIAKKGPDSDYAMKNNLKDISEGL 307

RESULT 4

A47186
 fimbrial adhesin fimA precursor - Streptococcus parasanguinis

C:Species: Streptococcus parasanguinis
 C:Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 17-Nov-2000
 C:Accession: A37186; S61912

R:Penno, J.C.; Leblanc, D.J.; Fives-Taylor, P.
 Infect. Immun. 57, 3527-3533, 1989

A:Title: Nucleotide sequence analysis of a type 1 fimbrial gene of Streptococcus sang
 A:Reference number: A37186; MUID:90035427; PMID:2572555

A:Accession: A37186
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-309 <FEN1>

A:Cross-references: GB:M26130; NID:g567768; PIDN:AAA53077.1; PID:g153834
 A:Experimental source: strain FW213

A:Note: the source is designated as Streptococcus sanguis
 R:Penno, J.C.; Shaikh, A.; Spatafora, G.; Fives-Taylor, P.
 Mol. Microbiol. 15, 849-863, 1995

A:Title: The fimA locus of Streptococcus parasanguis encodes an Afp-binding membrane
 A:Reference number: S61910; MUID:95319327; PMID:7596287

A:Accession: S61912
 A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-309 <FEN2>

A:Cross-references: EMBL:M26130; NID:g567768; PIDN:AAA53077.1; PID:g153834
 A:Experimental source: strain FW213

A:Note: the nucleotide sequence was submitted to the EMBL data library, October 1994
 A:Note: this publication is not cited in GenBank entry SRSYRA, release 117.0

A:Note: the source is designated as Streptococcus parasanguis
 C:Genetics:

A:Gene: fimA
 C:Superfamily: adhesin B

C:Keywords: blocked amino end; lipoprotein; membrane protein
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-309/Product: fimbrial adhesin fimA #status predicted <MAT>
 F:21/Binding site: sn-2,3-diacylglycerol (Cys) (covalent) #status predicted
 F:21/Modified site: fatty acylated amino end (Cys) (in mature form) #status predicted

Query Match 79.9%; Score 1178; DB 2; Length 309;
 Best Local Similarity 76.6%; Pred. No. 4.5e-72;
 Matches 219; Conservative 34; Mismatches 33; Indels 0; Gaps 0;

OY 2 STGAKTAKSDKLKVVATSIADMTAKAGDKIDLSIYVIGODPHEPEPLPDAKTSN 61
 DB 22 STGAKTAKSDKLKVVATSIADMTAKAGDKIDLSIYVIGODPHEPEPLPDAKTSN 81
 OY 62 ADVIFYNGINLEGGQAMFTKLVNAOKTKKNDYFAVSDGIDVYILEGASKEGDEPNA 121
 DB 82 ADVIFYNGINLEGGQAMFTKLVNAOKTKKNDYFAVSDGIDVYILEGASKEGDEPNA 141
 OY 122 LNLNGLIYSKNIKAKOLIAKPPKNEYKLVKAVAKLEKIDKESKDEKIPAEKKL 181
 DB 142 LNLNGLIYSKNIKAKOLIAKPPKNEYKLVKAVAKLEKIDKESKDEKIPAEKKL 201
 OY 182 VISEGCFYFSGKAVGSAYIWEINTEEGPDDISSLEKLYKISALFVSSVDRP 241
 DB 202 VISEGCFYFSGKAVGSAYIWEINTEEGPDDISSLEKLYKISALFVSSVDRP 261

C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AG1677
R:Glaser, P.; Frangouli, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Jernel, G.; Duchaud, E.; Durand, L.; Dussurgeat, O.; Eutlian, K.D.; Fsihi, H. D.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001
A:Authors: Kreier, J.; Kunz, M.; Kunst, F.; Kurupkat, G.; Madueno, E.; Maltournam, A.; Maok, C.; Schluerer, T.; Simoes, N.; Tlertzi, A.; Valquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1677
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-310 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC97191.1; PID:916414462; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lln1961
C:Superfamily: adhesin B

[illegible]

RESULT 9
 AG1305
 adhesion binding proteins and lipoproteins with multiple specificity for metal cations (C:Species: *Listeria monocytogenes*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 A:Accession: AG1305
 R:Diasser, P.; Frangeul, L.; Buchliesser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fshl, H. Science 294, 849-852, 2001
 A:Authors: Krefl, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
 A:Reference number: ABI077; MUID:21537279; PMID:11679669
 A:Accession: AG1305
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-310 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAC99925.1; PID:G16411301; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo1847
 C:Superfamily: adhesin B

Query Match	53.58;	Score 789;	DB 2;	Length 310;
Best Local Similarity	52.28;	Pred. NO. 6.7e-46;		
Matches 153;	Conservative 59;	Mismatches 71;	Indels 10;	Gaps 5;

```

QY      1 SSTGKTAKSD-KLKVATN5I1ADMTKALAGKDIDHSHVPIGDPHEPEPI,PEDEKRT 59
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      20 SSONDSKRTKGLNVATVYSILADIVKNGKNKIELHSHVPGVDPHEIDPLPANIQSA 79
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      60 SNADYIFXNGINLEDGQAMFTKLVKNAOKTR--NKDYFAVSDGIDVITYLEGASEKCK-- 115
        ::||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      80 ADADIDIFXNGINLT--GNGWEDRMLETADKRGREDKNOVELSKGVKRYL--TEKERTS 135
        ::||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      116 -EDPAMNTLENGITYSKNIAKOLIARPKRKEDYERKLYAKLEKLDEAKSKSDAI 174
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      136 ETDPMAMLDLHAGLITYTEVRDALYKADPPADADYKIEAKAKYIDKLTLQDEAKQPADL 195
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      175 AENKRLIYSEGCFFYFSKAYGVPSAYIWEINTEBEGTPDOISSLIEKLKIKYKSALEFVE 234
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      196 PENOKTIVLTSEGAFFYFARGLKALAYIWEINTESQGGPDMOKIOIVGVEKEKPNLFVE 255
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      235 SSVDRRPMETVSKDSGIPYSEIFPDSIACKKGPDSYAAKMNKLKISBGL 287
        ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      256 TSVDRSMHSYSEKTCVPIFAKIFTDSTAKGGEVDPTLEMMRYMLDKIHDLG 308

```

RESULT 10
 H89832
 hypothetical protein SA0567 [imported] - *Staphylococcus aureus* (strain N315)
 C|Species: *Staphylococcus aureus*
 C|Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C|Accession: H89832
 R|Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
 ma, A.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, K.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A|Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A|Reference number: A89758; MUID:21311952; PMID:11418146
 A|Accession: H89832
 A|Status: preliminary
 A|Molecule type: DNA
 A|Residues: 1-309 <NUP>
 A|Cross-references: GB:BA000018; PID:g13700522; PIDN:BAB41819.1; GSPDB:GN00149
 A|Experimental source: strain N315
 C|Genetics:
 A|Gene: SA0567
 C|Superfamily: adhesin B

[illegible]

RESULT 11
D83714
surface adhesin A precursor psaa [imported] - *Bacillus halodurans* (strain C-125)
C1Species: *Bacillus halodurans*

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: DB3714
 R:Yakami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: DB3714
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-300 <STO>
 A:Cross-references: GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BA04235.1; GSPDB:GN00
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: psaa
 C:Superfamily: adhesin B

Query Match 37.3%; Score 550; DB 2; Length 300;
 Best Local Similarity 40.4%; Pred. No. 7.7e-30;
 Matches 111; Conservative 61; Mismatches 93; Indels 10; Gaps 3;

QY 13 LKVVATNSIADMTKAIAGDKIDLSIVPGDPHEHEPELPEDAERTSNADYIFNGINL 72
 DB 34 LKIVTSFSLGVDNLNIGERSVYIVPGEPHEPEVPSDFQAVSDADYFVYNGL 93
 QY 73 EDGGQAMFTKLVKNAQTKNKDYFAVSDGIDVYILEGASERKEDPHAMNLNENGIYSK 132
 DB 94 EE---WQRLVEY---TSDVDVEVSPIDALPLE---ESDGLDPHAMDVKNAWKYVE 143
 QY 133 NIAKOLIAKDPKKNKTEYKNAKAYAKLEKDKKSKFDIAENKLIYVSEGCFFYS 192
 DB 144 VTRDLVLRDPDGAIIYVANAAYLQDLQLEEMWHDVTTTPERCRTIVISENARYRFG 203
 QY 193 KAYGVSAVYIEMINEEGTDPQISLIEKLVKIPSALEFVSSVDRMRPEVSKDSGIP 252
 DB 204 EAYGDTYVIGIMLNHEBETPEQISRYVDIYKELDIPAFVETLVKSMVTYSNDGVD 263
 QY 253 IYSEIFTDSIAKKGRGDSYVAMKWNLDKISEGL 287
 DB 264 IAGEVYTDVAVGLESGAETIYIDMKHNVDTFVSG 298

RESULT 12

664063
 Probable Mn transport protein - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 24-Sep-1999
 C:Accession: G64063; C41833
 R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
 . D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: G64063
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-293 <TTGR>
 A:Cross-references: GB:U32720; GB:142023; NID:g1573322; PIDN:AAC22021.1; PID:g1573330; T
 R:Kirkness, R.E.; Chong, P.; Klein, M.H.
 J. Bacteriol. 174, 2425-2430, 1992
 A:Title: Identification of two iron-repressed periplasmic proteins in *Haemophilus influe*
 A:Reference number: A41833; MUID:92210482; PMID:1556062
 A:Accession: C41833
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 23-42, 'D', 44-46 <HAR>
 C:Superfamily: adhesin B

Query Match 33.2%; Score 489.5; DB 2; Length 293;
 Best Local Similarity 34.5%; Pred. No. 8.8e-26;
 Matches 95; Conservative 62; Mismatches 111; Indels 7; Gaps 2;

QY 12 LKVVATNSIADMTKAIAGDKIDLSIVPGDPHEHEPELPEDAERTSNADYIFNGIN 71
 DB 23 KRVVYTFVVIDIDIAQNVAGNATVETISIRPAHEIHEPTKRDYKASADLILNGLN 82
 QY 72 LEDGGQAMFTKLVKNAQTKNKDYFAVSDGIDVYILEGASERKEDPHAMNLNENGIYS 131
 DB 83 LE---RMFERFEQN---VKDRPAVVYVTEGIDPLSTIEGCPYKADAPNPHAMSPSNALIYI 135
 QY 132 KNAKOLIAKDPKKNKTEYKNAKAYAKLEKDKKSKFDIAENKLIYVSEGCFFYS 191
 DB 136 ENIKNALVYDPONAVYERKNADYAKTKQDEPLRAKLAQIPEAKRLVYSEGAFFSYL 195
 QY 192 SKAYGVSAVYIEMINEEGTDPQISLIEKLVKIPSALEFVSSVDRMRPEVSKDSGI 251
 DB 196 AKDYNKEGYLMPINAEQGTPOQVAKYIDLVKKNIPVFESESTISAKPAQVAKESGA 255
 QY 252 PIYSEIFTDSIAKKGRGDSYVAMKWNLDKISEG 286
 DB 256 KYGCVLYVDSLAKNGPVPTIIDLNVVSTIVKG 290

RESULT 13

periplasmic-binding protein [imported] - *Yersinia pestis* (strain CO92)
 C:Species: *Yersinia pestis*
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
 C:Accession: AH0297
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.
 deno-Warraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrer
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11566360
 A:Accession: AH0297
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-323 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC91244.1; PID:g15980433; GSPDB:GN00175
 C:Genetics:
 A:Gene: yfeA
 C:Superfamily: adhesin B

Query Match 32.5%; Score 479.5; DB 2; Length 323;
 Best Local Similarity 35.0%; Pred. No. 4.7e-25;
 Matches 98; Conservative 60; Mismatches 113; Indels 9; Gaps 3;

QY 7 TAKSDKLKVVATNSIADMTKAIAGDKIDLSIVPGDPHEHEPELPEDAERTSNADYIF 66
 DB 50 TAK--KFKVYTFVVIDIDIAQNVAGNATVETISIRPAHEIHEPTKRDYKASADLIL 107
 QY 67 YGGINLEDGGQAMFTKLVKNAQTKNKDYFAVSDGIDVYILEGASERKEDPHAMNLNEN 126
 DB 108 NMGMINL---RMFEFEF---ESIDVPSAVVYAGTTPPIREGPYSGIANPHAMSPSN 160
 QY 127 GIYSKNIKQIADPKPKKTEYKNAKAYAKLEKDKKSKFDIAENKLIYVSEGS 186
 DB 161 ALIYIENIKALVENDPAEAETVYNNNAQYAKIRAKALDPLERLSRIYADBRMLVTSIG 220
 QY 187 CKRYSKAYGVSAVYIEMINEEGTDPQISLIEKLVKIPSALEFVSSVDRMRPEVYS 246
 DB 221 AFSYLAQDGFPEVYVIMPIINAEQGTPOQVRYVIDIRRNKIPVFESESTISDKPAQVYS 280
 QY 247 KDSGIPIYSEIFTDSIAKKGRGDSYVAMKWNLDKISEG 286
 DB 281 KETGAQYGVLYVDSLSEKGPVPTIISLNMVTDVIANG 320

RESULT 14

AC3106
 Hypoetical protein sita [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
 C:Accession: AC3106

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H.
erge, P.; Giller, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; Mclellan,
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, B.W.
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AC3106

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-299 <KUR>

A:Cross-references: GB:AE008669; PIDN:AAL45265.1; PID:917742950; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: sitA

A:Map position: linear chromosome

C:Superfamily: adhesin B

Query Match 32.2%; Score 475.5; DB 2; Length 299;

Best Local Similarity 33.2%; Pred. No. 7.9e-25;

Matches 92; Conservative 64; Mismatches 114; Indels 7; Gaps 2;

QY 11 DKLKVAATNSIIADMTALAGDKIDLHSIYPIGDDPHEPEPLPDAEKTSNADYIFYNGI 70

DB 28 EKPKVVTFTTIIADMARNVAGDADVESITKPGAIEHNYOPTPRDILKARKADLVLRNGL 87

QY 71 NLEDGGQAMFTKLVKNQKTKNKDYFAVSDGIDVYILEGASEKGEDPHAMNLNENGIY 130

DB 88 NLE----LWFEKFLANISGVDP---VTVSDGVOPMAISGAYQCKPNNHAMSPDNALIT 140

QY 131 SKNIKQILAKDPKPKETTERKLNKAYAKLEKLDREKSKEDAIENKKLIYVSEGCERY 190

DB 141 VENIRKGLAEIDPAHADVYTANAKAYSDKIKATVQPIRDALSVLPDKNRMLVTVSEGAFSY 200

QY 191 FSKAYGPSAYIWEINTEEGTPOQISLIEKLVKPSALFVSSVDRRPMETVSKDSG 250

DB 201 LARDFGLKEFLMPVNVADSGTPOQVGVVIDAMREHNIQVIFSESTVSADPAKQVAKETG 260

QY 251 IPIYSEIFTSIAKKRPGDSYVYAMKMNLDKISEGL 287

DB 261 ASYGILYVDLSLSEADGPVPYIIDLCLVTSSETIAKGL 297

RESULT 15

A96181

sitA protein (AF128999) [imported] - *Agrobacterium tumefaciens* (strain C58, Cereon)

C:Species: *Agrobacterium tumefaciens*

C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002

C:Accession: A96181

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*

A:Reference number: A97359; PMID:11743194

A:Accession: A96181

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-300 <KUR>

A:Cross-references: GB:AE007870; PIDN:AAK8971.1; PID:915158753; GSPDB:GN00170

C:Genetics:

A:Gene: AGR_L_798

A:Map position: linear chromosome

C:Superfamily: adhesin B

Query Match 32.2%; Score 475.5; DB 2; Length 300;

Best Local Similarity 33.2%; Pred. No. 7.9e-25;

Matches 92; Conservative 64; Mismatches 114; Indels 7; Gaps 2;

QY 11 DKLKVAATNSIIADMTALAGDKIDLHSIYPIGDDPHEPEPLPDAEKTSNADYIFYNGI 70

DB 29 EKPKVVTFTTIIADMARNVAGDADVESITKPGAIEHNYOPTPRDILKARKADLVLRNGL 88

Search completed: May 21, 2003, 10:59:25
Job time : 47 secs

QY 71 NLEDGGQAMFTKLVKNQKTKNKDYFAVSDGIDVYILEGASEKGEDPHAMNLNENGIY 130

DB 89 NLE----LWFEKFLANISGVDP---VTVSDGVOPMAISGAYQCKPNNHAMSPDNALIT 141

QY 131 SKNIKQILAKDPKPKETTERKLNKAYAKLEKLDREKSKEDAIENKKLIYVSEGCERY 190

DB 142 VENIRKGLAEIDPAHADVYTANAKAYSDKIKATVQPIRDALSVLPDKNRMLVTVSEGAFSY 201

QY 191 FSKAYGPSAYIWEINTEEGTPOQISLIEKLVKPSALFVSSVDRRPMETVSKDSG 250

DB 202 LARDFGLKEFLMPVNVADSGTPOQVGVVIDAMREHNIQVIFSESTVSADPAKQVAKETG 261

QY 251 IPIYSEIFTSIAKKRPGDSYVYAMKMNLDKISEGL 287

DB 262 ASYGILYVDLSLSEADGPVPYIIDLCLVTSSETIAKGL 298